

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2005, 23:34:02 ; Search time 88.5 Seconds
(without alignments)
2459.137 Million cell updates/sec

Title: US-10-728-947-2

Perfect score: 2155

Sequence: 1 VAEIMHVFAREILDSRGNPT.....QLLGDAGVVAGRSAPFRFQG 425

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

UniProt_03:*

1: uniprot_sprot:*

2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2152	99.9	425	1 ENO_CORGL	Q8nr61 corynebacte
2	2029	94.2	425	1 ENO_COREF	Q8fg97 corynebacte
3	1917	89.0	425	2 Q6NI61	Q6ni61 corynebacte
4	1599.5	74.2	426	1 ENO1_STRCO	Q8f2d3 streptomyce
5	1371.5	72.9	426	2 Q6ADR6	Q6adr6 leifsonia x
6	1558.5	72.3	429	2 Q741U7	Q741u7 mycobacteri
7	1554.5	72.1	428	2 Q82HH5	Q82hh5 streptomyce
8	1551.5	72.0	429	1 ENO_MYCTU	P96377 mycobacteri
9	1551.5	72.0	429	2 Q7U0U6	Q7u0u6 mycobacteri
10	1508.5	70.0	426	2 Q6AAB8	Q6aab8 propionibac
11	1500.5	69.6	429	1 ENO_MYCLE	Q9cd42 mycobacteri
12	1436	66.6	428	2 Q83H73	Q83h73 tropheryma
13	1436	66.6	446	2 Q83PF7	Q83ff7 tropheryma
14	1418	65.8	424	1 ENO_AGRT5	Q8ufh1 agrobacteri
15	1412	65.5	429	1 ENO_THETN	Q8r967 thermoanaer
16	1408.5	65.4	426	1 ENO_CAUCR	Q8a7j9 caulobacter
17	1406	65.2	429	1 ENO_BACHD	Q9k7i7 bacillus ha
18	1404.5	65.2	429	2 Q84FY9	Q84fy9 methylobact
19	1402	65.1	425	1 ENO_BRUSU	Q8xg03 brucella su
20	1400	65.0	424	1 ENO_RHILO	Q98m23 rhizobium l
21	1395	64.7	425	1 ENO_BRUME	Q8vfh0 rhizocella m
22	1394.5	64.7	430	2 Q65EN2	Q65en2 bacillus li
23	1393	64.6	424	1 ENO_RHIME	Q92q98 rhizobium m
24	1381.5	64.1	430	1 ENO_BACSU	P37869 bacillus su
25	1380.5	64.1	431	1 ENO_CLOPE	Q8xku4 clostridium
26	1379	64.0	431	2 Q815K8	Q815k8 bacillus ce
27	1377	63.9	431	2 Q81X78	Q81x78 bacillus an
28	1377	63.9	431	2 Q6HBF3	Q6hbf3 bacillus th
29	1375	63.8	431	2 Q63IM2	Q63lm2 bacillus ce
30	1374	63.8	428	2 Q74AR6	Q74ar6 geobacter s
31	1373	63.7	432	2 Q7VDY0	Q7vdy0 prochloroco

32 1370.5 63.6 427 1 ENO_RALSO
33 1366.5 63.4 430 2 Q7U3T1
34 1366 63.4 431 2 Q72XV5
35 1365.5 63.4 430 2 Q67SV9
36 1360.5 63.1 430 1 ENO_LISIN
37 1360.5 63.1 430 1 ENO_LISMO
38 1360.5 63.1 431 2 Q71WX1
39 1360.5 63.1 431 2 Q898R0
40 1359.5 63.1 429 1 ENO_OCEIH
41 1359.5 63.1 434 2 Q8GDZ5
42 1358.5 63.0 431 1 ENO_CLOAB
43 1353 62.8 430 2 Q7V377
44 1351 62.7 431 2 Q7V483
45 1346 62.5 432 1 ENO_ENTFA

ALIGNMENTS

RESULT 1

ID ENO_CORGL STANDARD; PRT; 425 AA.
AC Q8NR61;
DT 28-FEB-2003 (Rel. 41, Last Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Enolase (EC 4.2.1.11) (2-phosphoglycerate dehydratase) (2-phospho-D-
glycerate hydro-lyase).
DE glycerate hydro-lyase).
GN Name=eno; OrderedLocusNames=Cgl0974, cgl111;
OS Corynebacterium glutamicum (Brevibacterium flavum).
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=1718;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
RA Nakagawa S.;
RT "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
RX PubMed=12948626; DOI=10.1016/S0168-1656(03)00154-8;
RA Kalinowski J., Bathe B., Bartels D., Bischoff N., Bott M.,
Burkovski A., Dusch N., Eggeling L., Eikmanns B.J., Gaigalat L.,
Goesmann A., Hartmann M., Huthmacher K., Kraemer R., Linke B.,
McHardy A.C., Meyer F., Moeckel B., Pfeifferle W., Puhler A.,
Rey D.A., Rueckert C., Rupp O., Sahn H., Wendisch V.F., Wiesgraebe I.,
Tauch A.;
RT "The complete Corynebacterium glutamicum ATCC 13032 genome sequence
and its impact on the production of L-aspartate-derived amino acids
and vitamins.";
RL J. Biotechnol. 104:5-25 (2003).
CC -1- CATALYTIC ACTIVITY: 2-phospho-D-glycerate = phosphoenolpyruvate +
H(2O).
CC -1- COFACTOR: Magnesium is required for catalysis and for stabilizing
the dimer (By similarity).
CC -1- PATHWAY: Glycolysis.
CC -1- SUBUNIT: Homodimer (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the enolase family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AP005277; BAB98367.1; -.
DR EMBL; BX927150; CAF19681.1; -.
DR HSSP; P08324; 1E9I.

[illegible]

DR	HMAP; MF 00318; -: 1.
DR	InterPro; IPR000941; Enolase.
DR	Pfam; PF00113; Enolase_C; 1.
DR	Pfam; PF03952; Enolase_N; 1.
DR	PRINTS; PR00148; ENOLASE.
DR	ProDom; PD00902; Enolase; 1.
DR	TIGRFAMs; TIGR01060; eno; 1.
DR	KW PROSITE; PS00164; ENOLASE; 1.
KW	Complete proteome; Glycolysis; Lyase; Magnesium.
FT	ACT_SITE 154 154 By similarity.
FT	METAL 241 241 Magnesium (By similarity).
FT	METAL 282 282 Magnesium (By similarity).
FT	METAL 309 309 Magnesium (By similarity).
SQ	SEQUENCE 425 AA; 44949 MW; B7AABD82C165FD1D CRC64;

Query Match	99.9%; Score 2152; DB 1; Length 425;
Best Local Similarity	99.8%; Pred. No. 4.1e-119;
Matches 424; Conservative 1; Mismatches 0; Indels 0; Gaps 0;	

Qy	1	VAEIMHVFAREILDSRCGNPTVEAEVFLDDGSHGVAGVPVGASTGVHEAHRLDGGDRYL	60
Dd	1	MAEIMHVFAREILDSRCGNPTVEAEVFLDDGSHGVAGVPVGASTGVHEAHRLDGGDRYL	60
Qy	61	KGVLKAVENNEETIGDELAGEADDQRLLIDEAMIKLDTGTANKSRIGANAILGVSMVAKA	120
Dd	61	KGVLKAVENNEETIGDELAGEADDQRLLIDEAMIKLDTGTANKSRIGANAILGVSMVAKA	120
Qy	121	AADSAGLPFRYIIGPNNAHVLVPVMNIIINGGAHADSGVDVOEPWIPIGAIETSEALRN	180
Dd	121	AADSAGLPFRYIIGPNNAHVLVPVMNIIINGGAHADSGVDVOEPWIPIGAIETSEALRN	180
Qy	181	GAEVYHALKSVIKKGLSTGLDGEGFPAPSVGSTREALDLIVEATEKAAGTFPGKDIALAL	240
Dd	181	GAEVYHALKSVIKKGLSTGLDGEGFPAPSVGSTREALDLIVEATEKAAGTFPGKDIALAL	240
Qy	241	DVASSEFFKCGTYHFEGQHSAEAEMANYVAELVDAYPIVSIEDPLQEDDWGYTNLTATI	300
Dd	241	DVASSEFFKCGTYHFEGQHSAEAEMANYVAELVDAYPIVSIEDPLQEDDWGYTNLTATI	300
Qy	301	GDKVQIVGDGDFVTNNPERLKKEGIKAANSILVKVNIQTTLTETFDVDMHAHRAGYTSM	360
Dd	301	GDKVQIVGDGDFVTNNPERLKKEGIKAANSILVKVNIQTTLTETFDVDMHAHRAGYTSM	360
Qy	361	SHRSGETEDTTIADLAVALNCQIKTGAPASRDRAKYNQLRIEQLLGDAGVYAGRSAF	420
Dd	361	SHRSGETEDTTIADLAVALNCQIKTGAPASRDRAKYNQLRIEQLLGDAGVYAGRSAF	420
Qy	421	PRFQG 425	
Dd	421	PRFQG 425	

RESULT 2	
ENO_COREF	
ID	ENO_COREF STANDARD; PRT; 425 AA.
AC	O8FGS7:
DT	10-OCT-2003 (Rel. 42, Created)
DT	10-OCT-2003 (Rel. 42, Last sequence update)
DT	05-JUL-2004 (Rel. 44, Last annotation update)
DE	Enolase (EC 4.2.1.11) (2-phosphoglycerate dehydratase) (2-phospho-D-glycerate hydro-lyase).
DE	Name=eno; OrderedLocusNames=Cel1042;
GN	Corynebacterium efficiens.
OS	Corynebacteria; Actinobacteridae; Actinomycetales;
OC	Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX	NCBI_Taxid=152794;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=Y-314 / AJ 12310 / DSM 44549 / JCM 11189;
RX	MEDLINE=22723752; PubMed=12840036; DOI=10.1101/gr.1285603;
RA	Nishio Y., Nakamura Y., Kawarabayashi Y., Usuda Y., Kimura E., Sugimoto S., Matsui K., Yamagishi A., Kikuchi H., Ikeo K., Gojobori T.;

```

RESULT 3
O6NI61 PRELIMINARY; PRT; 425 AA.
AC Q6NI61;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Enolase (EC 4.2.1.11).
GN Name=eno; OrderedLocusNames=DIP0917;
OS Corynebacterium diphtheriae.
OC Bacteria; Actinobacteridae; Actinobacteriales; Actinomycetales;
OC Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=1717;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Biotype gravis / NCTC 13129;
RX MEDLINE=22965443; PubMed=14602910; DOI=10.1093/nar/gkg874;
RA Cerdeno-Tarraga A.-M., Efstathiou A., Dover L.G., Holden M.T.G.,
RA Pallen M.J., Bentley S.D., Beara G.S., Churcher C.M., James K.D.,
RA De Zuyasa A., Chillingworth T., Cronin A., Dowd L., Feltwell T.,
RA Hamlin N., Holroyd S., Jagels K., Moule S., Quail M.A.,
RA Rabinowitsch E., Rutherford K.M., Thomson N.R., Unwin L.,
RA Whitehead S., Barrell B.G., Parkhill J.;
RA "The complete genome sequence and analysis of Corynebacterium
RT diphtheriae NCTC13129."
RL Nucleic Acids Res. 31:6516-6523 (2003).
CC -!- CATALYTIC ACTIVITY: 2-phospho-D-glycerate = phosphoenolpyruvate +
CC H(2)O.
CC -!- COFACTOR: Magnesium is required for catalysis and for stabilizing
CC the dimer (By similarity).
CC -!- PATHWAY: Glycolysis.
CC -!- SUBUNIT: Homodimer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the enolase family.
CC EMBL: BX248356; CAE9433.1; -.
DR GO: GO:0000015; C:phosphopyruvate hydratase complex; IEA.
DR GO: GO:0016829; F:lyase activity; IEA.
DR GO: GO:0000287; F:magnesium ion binding; IEA.
DR GO: GO:0004634; F:phosphopyruvate hydratase activity; IEA.
DR GO: GO:0006096; F:glycolysis; IEA.
DR InterPro: IPR000941; Enolase.
DR Pfam: PF00113; Enolase_C; 1.
DR Pfam: PF03952; Enolase_N; 1.
DR PRINTS: PR00148; ENOLASE.
DR ProDom: PD000902; Enolase; 1.
DR TIGRFAMs: TIGR01060; eno; 1.
DR PROSITE: PS00164; ENOLASE; 1.
DR Complete proteome: Glycolysis; Lyase; Magnesium.
SQ SEQUENCE 425 AA; 45029 MW; 858F477EB7ECDCA6 CRC64;

Query Match 89.0%; Score 1917; DB 2; Length 425;
Best Local Similarity 88.0%; Pred. No. 3.1e-105;
Matches 374; Conservative 22; Mismatches 29; Indels 0; Gaps 0;

Qy 1 VAETIMHVFARILDSRGNTPTVEAEVFLDDGSHGVAGVPSTGASTGVHEAHELRDGGERYLG 60
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 MADIMHVFARILDSRGNTPTVEAEVFLDDGSHGVAGVPSTGASTGVHEAHELRDGGERYLG 60
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 61 KGVLLKAVENNEETIGDELAGEADDQRLIDAMTKLDGTANKSRILGANAILGVSMVAKA 120
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 KGVLLNAVNNVNEETADATAGAEADDQRLIDQAMTALDGTENKSRILGANAILGVSIYAKA 120
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 121 AADSAGLPLFRYIGGNPAHVLVPVMNIIINGGAHADSGVDVQEPMPAPIGAETSEALRN 180
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 121 AAESAGLPLFRYIGGNPAHVLVPVMNIIINGGAHADSGVDVQEPMPAPIGAETSEALRM 180
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 181 GAETHYALKSVIKKGLSTGLDGGGFPSPVSTREALDLIVEAEKAGETPGKDIALAL 240
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 181 GAETHYLSKSVIKKGLSTGLDGGGFPSPVESTKAALDLIVEAEKAGFKPGADIALAL 240
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 241 DVASSEFFKDTYHFEGGQHSAAEMANVYAEVLDAVPVTSIEDPLQEDDMEGYTALTAI 300
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

KW Complete proteome; Glycolysis; Lyase; Magnesium.
FT ACT SITE 155 155 By similarity.
FT METAL 242 242 Magnesium (By similarity).
FT METAL 283 283 Magnesium (By similarity).
FT METAL 310 310 Magnesium (By similarity).
SQ SEQUENCE 426 AA; 45483 MW; 63AED7B7A4E15001 CRC64;

Query Match 74.2%; Score 1599.5; DB 1; Length 426;
Best Local Similarity 74.8%; Pred. No. 1.7e-85;
Matches 314; Conservative 46; Mismatches 59; Indels 1; Gaps 1;

QY 7 VFAREILDSRGNTVEAEVFLDDGSHGVAGVPSGASTGVHAEHLRDGG-DRYLGGKVLK 65
DB 7 VFAREILDSRGNTVEAEVFLDDGSHGVAGVPSGASTGVHAEHLRDGG-DRYLGGKVLK 65
QY 7 WAREILDSRGNTVEAEVFLDDGSHGVAGVPSGASTGVHAEHLRDGG-DRYLGGKVLK 66
DB 7 WAREILDSRGNTVEAEVFLDDGSHGVAGVPSGASTGVHAEHLRDGG-DRYLGGKVLK 66
QY 66 AVENNEEIGDELADQRLIDEAMIKLDTGANKSRGLGANAILGVSMVAKAADS 125
DB 66 AVENNEEIGDELADQRLIDEAMIKLDTGANKSRGLGANAILGVSMVAKAADS 125
QY 67 AVLAVIEIQIGPELVGYDATERLIDQAMFDLDTDNKSGSLGANAILGVSLAVAHAASEAS 126
DB 67 AVLAVIEIQIGPELVGYDATERLIDQAMFDLDTDNKSGSLGANAILGVSLAVAHAASEAS 126
QY 126 GLPLFRYIGGNPAHVLVPMNMNIINGGAHDSGVVDVQEFMIAPIGATTFSEALRNGAEVY 185
DB 126 GLPLFRYIGGNPAHVLVPMNMNIINGGAHDSGVVDVQEFMIAPIGATTFSEALRNGAEVY 185
QY 127 DLPLFRYIGGNPAHVLVPMNMNIINGGAHDSGVVDVQEFMIAPIGATTFSEALRNGAEVY 186
DB 127 DLPLFRYIGGNPAHVLVPMNMNIINGGAHDSGVVDVQEFMIAPIGATTFSEALRNGAEVY 186
QY 186 HALSKVIXEKLSTGLGDEGFPAPSVGSTREALDLIVEATEKAGFTPGKDIALADVASS 245
DB 186 HALSKVIXEKLSTGLGDEGFPAPSVGSTREALDLIVEATEKAGFTPGKDIALADVASS 245
QY 187 HTLKKVLNKGATLGLGDEGFPAPSVGSTREALDLIVEATEKAGFTPGKDIALADVASS 246
DB 187 HTLKKVLNKGATLGLGDEGFPAPSVGSTREALDLIVEATEKAGFTPGKDIALADVASS 246
QY 246 EFFKDGTYHPEGGQHSAAEMANVYAEIVDAIPVSIEDPLQEDDWEYTNLTATIGDKVQ 305
DB 246 EFFKDGTYHPEGGQHSAAEMANVYAEIVDAIPVSIEDPLQEDDWEYTNLTATIGDKVQ 305
QY 247 EFKDGTYHPEGGQHSAAEMANVYAEIVDAIPVSIEDPLQEDDWEYTNLTATIGDKVQ 306
DB 247 EFKDGTYHPEGGQHSAAEMANVYAEIVDAIPVSIEDPLQEDDWEYTNLTATIGDKVQ 306
QY 306 IVDGDFVTNPERLKEGIAKKAANSILVKVNIQITLTETFDVDMAHRAGYTSMHSRSG 365
DB 306 IVDGDFVTNPERLKEGIAKKAANSILVKVNIQITLTETFDVDMAHRAGYTSMHSRSG 365
QY 307 LVGDGLFVTNPERLARGTEENSANALLVKVNIQISLTETTDVAVLAQNGFKCMHSRSG 366
DB 307 LVGDGLFVTNPERLARGTEENSANALLVKVNIQISLTETTDVAVLAQNGFKCMHSRSG 366
QY 366 ETEDTTADLAVANLCOQIKTGAPARSDRVAKYNQLLRIEQLLDAGVYAGRSAPFRFQG 425
DB 366 ETEDTTADLAVANLCOQIKTGAPARSDRVAKYNQLLRIEQLLDAGVYAGRSAPFRFQG 425
QY 367 ETEDTTADLAVANLCOQIKTGAPARSDRVAKYNQLLRIEQLLDAGVYAGRSAPFRFQG 426
DB 367 ETEDTTADLAVANLCOQIKTGAPARSDRVAKYNQLLRIEQLLDAGVYAGRSAPFRFQG 426

RESULT 5
Q6ADR6 PRELIMINARY; PRT; 426 AA.
AC Q6ADR6;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Enolase.
GN Name=eno; OrderedLocusNames=Lxx17200;
OS Leifsonia xyl (subsp. xyl).
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Micrococciaceae; Microbacteriaceae; Leifsonia.
OX NCBI_TaxID=59736;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CTCB07;
RX PubMed=15305603;
RA Monteiro-Vitorello C.B., Camargo L.E.A., Van Sluys M.A.,
RA Kitajima J.P., Truffi D., do Amaral A.M., Harakava R.,
RA de Oliveira J.C.F., Wood D., de Oliveira M.C., Miyaki C.Y.,
RA Takita M.A., da Silva A.C.R., Furlan L.R., Carraro D.M., Camarotte G.,
RA Almeida N.F. Jr., Carrer H., Coutinho L.L., El-Dorzy H.A.,
RA Ferro M.I.T., Gagliardi P.R., Gigliotti E., Goldman M.H.S.,
RA Goldman G.H., Kimura E.T., Ferro E.S., Kuramae E.E., Lemos E.G.M.,
RA Lemos M.V.F., Mauro S.M.Z., Machado M.A., Marino C.L., Menck C.F.,
RA Nunes L.R., Oliveira R.C., Pereira G.G., Siqueira W., de Souza A.A.,
RA Tsai S.M., Zanca A.S., Simpson A.J.G., Brumbley S.M., Setubal J.C.;
RT "The genome sequence of the Gram-positive sugarcane pathogen Leifsonia
xyl subsp. xyl.";
RL Mol. Plant Microbe Interact. 17:897-936(2004).
CC -I- CATALYTIC ACTIVITY: 2-phospho-D-glycerate = phosphoenolpyruvate +
CC H(2O).
CC -I- COPACTOR: Magnesium is required for catalysis and for stabilizing

```

```

CC the dimer (By similarity).
CC -I- PATHWAY: Glycolysis.
CC -I- SUBUNIT: Homodimer (By similarity).
CC -I- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -I- SIMILARITY: Belongs to the enolase family.
CC EMBL; A5016822; AAT89480.1; -.
DR GO; GO:0000015; C:phosphopyruvate hydratase complex; IEA.
DR GO; GO:0004634; F:phosphopyruvate hydratase activity; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR000941; Enolase.
DR Pfam; PF00113; Enolase C; 1.
DR PRINTS; PR00148; ENOLASE.
DR ProDom; PD000902; Enolase; 1.
DR TIGRFAMs; TIGR01060; eno; 1.
DR PROSITE; PS00164; ENOLASE; 1.
DR Complete proteome; Glycolysis; Lyase; Magnesium.
SQ SEQUENCE 426 AA; 45227 MW; 43DE1F7B691BC2E1 CRC64;

Query Match 72.9%; Score 1571.5; DB 2; Length 426;
Best Local Similarity 73.0%; Pred. No. 7.6e-85;
Matches 311; Conservative 46; Mismatches 68; Indels 1; Gaps 1;

QY 1 VAEIMHVFAREILDSRGNTVEAEVFLDDGSHGVAGVPSGASTGVHAEHLRDGG-DRYL 59
DB 1 MAEIAVGAAREILDSRGNTVEAEVFLDDGSHGVAGVPSGASTGVHAEHLRDGG-DRYL 60
QY 60 GKVLAENVNEEIGDELADQRLIDEAMIKLDTGANKSRGLGANAILGVSMVAK 119
DB 61 GKVLAENVNEEIGDELADQRLIDEAMIKLDTGANKSRGLGANAILGVSMVAK 120
QY 120 AAADSAGLPLFRYIGGNPAHVLVPMNMNIINGGAHDSGVVDVQEFMIAPIGATTFSEALR 179
DB 121 AAADSAGLPLFRYIGGNPAHVLVPMNMNIINGGAHDSGVVDVQEFMIAPIGATTFSEALR 180
QY 180 NGAEVYHALKSVIXEKLSTGLGDEGFPAPSVGSTREALDLIVEATEKAGFTPGKDIALA 239
DB 181 WGVETVYSLKALLKSKGLNTGLGDEGFPAPSVGSTREALDLIVEATEKAGFTPGKDIALA 240
QY 240 LVASSEFFKDGTYHPEGGQHSAAEMANVYAEIVDAIPVSIEDPLQEDDWEYTNLTAT 299
DB 241 LDVASTEFFGVYRFGDQRTAAENSAIYVTDLANNYPVLSIEDPLAEDDWEYTNLTAT 300
QY 300 IGDVQVIGDGFVTNPERLKEGIAKKAANSILVKVNIQITLTETFDVDMAHRAGYTSM 359
DB 301 IGSTLQVIGDGLFVTNPERLKEGIAKKAANSILVKVNIQITLTETFDVDMAHRAGYTSM 360
QY 360 MSHRSGETEDTTIADLAVANLCOQIKTGAPARSDRVAKYNQLLRIEQLLDAGVYAGRS 419
DB 361 LSHRSGETEDTTIADLAVANLCOQIKTGAPARSDRVAKYNQLLRIEQLLDAGVYAGRS 420
QY 420 FPRFQG 425
DB 421 FPRFQG 426

RESULT 6
Q741U7 PRELIMINARY; PRT; 429 AA.
AC Q741U7;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Eno.
GN Name=eno; OrderedLocusNames=MAP0990;
OS Mycobacterium paratuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1770;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=k10;
RA Li L., Bannantine J., Zhang Q., Amonsin A., Alt D., Kapur V.;

```


	[1]	SEQUENCE FROM N.A.	
RN	RP	STRAIN=MA-4680;	
RC	RC	MEDLINE=22698306; PubMed=11572948; DOI=10.1073/pnas.211433198;	
RX	MX	Ikedo H., Ishikawa J., Hanamoto A., Takahashi C.,	
RA	OMura S., Ikeda H., Shinoe M., Takahaehi Y., Horikawa H., Nakazawa H., Osonoe T.,		
RA	Kikuchi H., Shiba T., Sakaki Y., Hattori M.;		
RA	"Genome sequence of an industrial microorganism Streptomyces		
RT	avermittilis: deducing the ability of producing secondary		
RT	metabolites.";		
RL	Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).		
RN	[2]	SEQUENCE FROM N.A.	
RP	RP	STRAIN=MA-4680;	
RC	RC	MEDLINE=22698306; PubMed=12692562;	
RX	IX	Ikedo H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,	
RA	Sakaki Y., Hattori M., Omura S.;		
RA	"Complete genome sequence and comparative analysis of the industrial		
RT	microorganism Streptomyces avermittilis.";		
RT	Microorganism Streptomyces avermittilis;"		
RL	Nat. Biotechnol. 21:526-531(2003).		
CC	-I- CATALYTIC ACTIVITY: 2-phospho-D-glycerate = phosphoenolpyruvate +		
CC	H ₂ O.		
CC	-I- COFACTOR: Magnesium is required for catalysis and for stabilizing		
CC	the dimer (By similarity).		
CC	-I- PATHWAY: Glycolysis.		
CC	-I- SUBUNIT: Homodimer (By similarity).		
CC	-I- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).		
CC	-I- SIMILARITY: Belongs to the enolase family.		
DR	EMBL; AP005035; BAC71245.1; -		
DR	HSP; Q9NDH8; 10EP.		
DR	GO: GO:0000015; C:phosphopyruvate hydratase complex; IEA.		
DR	GO: GO:0016829; F:lyase activity; IEA.		
DR	GO: GO:0000287; F:magnesium ion binding; IEA.		
DR	GO: GO:0004634; F:phosphopyruvate hydratase activity; IEA.		
DR	GO: GO:0006096; P:glycolysis; IEA.		
DR	InterPro: IPR000941; Enolase.		
DR	Pfam: PF00113; Enolase_C; 1.		
DR	Pfam: PF03952; Enolase_N; 1.		
DR	PRINTS; PR00148; ENOLASE.		
DR	ProDom: PD000902; Enolase; 1.		
DR	TIGRFAMS; TIGR01060; eno; 1.		
DR	PROSITE; PS00164; ENOLASE; 1.		
KW	Complete proteome; Glycolysis; Lyase; Magnesium.		
SQ	SEQUENCE 428 AA; 45869 MW; 09SA3976FC5F70D CRC64;		
	Query Match 72.1%; Score 1554.5; DB 2; Length 428;		
	Best Local Similarity 73.7%; Pred. No. 7.7e-84;		
	Matches 309; Conservative 42; Mismatches 67; Indels 1; Gaps 1		
Qy	7 VFAREILDSRGNPVTVEAEVFLDDGSHGVAGVPSGASTGVHEAHLRDGG-DRYLKGKVLK 65		
Dd			
Dd	7 VWAREILDSRGNPVTVEEVGLDDGTGRAAVPSGASTGAFAEIELRDGDPNRYQGKVGK 66		
Qy	66 AVENVNEIIGDELAGEADDOILLDEAMIKLDGTANKSRGLCANAILGVSMVAKAADA 125		
Dd			
Dd	67 AVLAVIEIQIPELGVGYATEQRLLDQAFMDATDNKSLGANATLGVS LAVAHAASEAS 126		
Qy	126 GLPLFRYIGPNAHVLPVPMNIINGGAHADSGVDVDFMFIPTICAETFSFSEALRNGAEVY 185		
Dd			
Dd	127 DLPLFRYIGPNAHVLPVPMNIINGGSHADSNDVIDIQEFMIPIGAESFSEALRNGAEVY 186		
Qy	186 HALKSIVKEKGLSTGLDGEGFAPSVGSTREALLDIVEAIKAGFTPGKDIALALDVASS 245		
Dd			
Dd	187 HTLKVKLVTKGLSTGLDGEGFAPNLSENRAALLDLIEAKQAGYIPGEIQIALALDVAA 246		
Qy	246 EFFKDGTYHFEGGQHSAAEMANVYAELVDAYPIVISIEDPLOQDDWEGVTNLTATGDKVK 305		
Dd			
Dd	247 EFKDGKYFEFGKSRSAEMTEYYEELVSAYPLVSIEDPLVEDDWAGWKVITDKLGDKVK 306		
Qy	306 IVGDDFFVTTPERLKEGIKAANSILVKVNOIGTLTTFDAVDMAHPAGYTSMMHSRG 365		
Dd			
Dd	307 IVGDDFFVTTPERLARGIEGSANALLVKVNOIGLSLTETLDAVELAQNRGFKCMMHSRG 366		

CC -!- CATALYTIC ACTIVITY: 2-phospho-D-glycerate = phosphoenolpyruvate + H2O.

CC -!- COFACTOR: Magnesium is required for catalysis and for stabilizing the dimer (By similarity).

CC -!- PATHWAY: Glycolysis.

CC -!- SUBUNIT: Homodimer (By similarity).

CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).

CC -!- SIMILARITY: Belongs to the enolase family.

DR EMBL; BX248337; CAD93911.1; -.

DR HSSP; P08324; 1B91.

DR GO; GO:0000015; C:phosphopyruvate hydratase complex; IEA.

DR GO; GO:0016829; F:lyase activity; IEA.

DR GO; GO:0000287; F:magnesium ion binding; IEA.

DR GO; GO:0004634; F:phosphopyruvate hydratase activity; IEA.

DR GO; GO:0006036; F:glycolysis; IEA.

DR InterPro; IPR000941; Enolase.

DR Pfam; PF00113; Enolase_C; 1.

DR Pfam; PF03952; Enolase_N; 1.

DR ProDom; PD000902; Enolase; 1.

DR TIGRFAMs; TIGR01060; eno; 1.

DR DR PROSITE; PS00164; ENOLASE; 1.

DR Complete proteome; Glycolysis; Lyase; Magnesium.

KW SEQUENCE 429 AA; 4862 MW; F808PB8E0D891C9A CRC64;

Query Match 72.0%; Score 1551.5; DB 2; Length 429;
 Best Local Similarity 73.2%; Pred. No. 1.2e-83;
 Matches 308; Conservative 41; Mismatches 71; Indels 1; Gaps 1;

QY 4 IMHVFAREILDSRGNPTVEAEVFLDDGSHGVAGVPSGASTGVHAEHLRDCGDRYLGKV 63
 DB 4 IEQVGAEREILDSRGNPTVEEVALIDGTFARAAVPSGASTGEHBEAVELRDCGDRYGGKV 63

QY 64 LKAVENNEEITGDELAGLEADDQRLIDEMTKLDGTANKSRIGANAILGVSMVAKAAAD 123
 DB 64 QKAVQAVLDEIGPAVIGLNADDQRLVDAQLVLDGTPDKSLGNAILGVSLAVAKAAAD 123

QY 124 SAGLPLFRYIGGNNAHLFVPMNMIINGGAHDSGVVQEFMIAPIGAETFSEALRNGAE 183
 DB 124 SAEELPLFRYVGGPNAHLFVPMNNILNGCAHADTAVIDIQEFWVAPIGAPSFEALRWGAE 183

QY 184 VYHAKSVIKEGLSTGLDGGGAPSVGSTREALDLIVEAIEKAGFTPGKDIALALDVA 243
 DB 184 VYHAKSVLKEGLSTGLDGGGAPVAGVTAAALDLISRAIESAGLRPGADVALALDAA 243

QY 244 SSEFFPKDGT-YHFGGQGHSAEMANVYAEVDYPIVSIEDPLEDDWEGYTNLTATIGD 302
 DB 244 ATEFTDGTGVFGGTRTADQMTFAYAGILGNPLVSIEDPLEDDWDGHAALTATIGD 303

QY 303 KVQIVGDDFVFTNPERLKEGIKAAANSILVKVNIQIGTLTETFDVADMAHRAGYTSMMSH 362
 DB 304 RVQIVGDDIFVTNPERLEEGERGVANALLVKVNIQIGTLTETDAVTLAHHGGYRTMISH 363

QY 363 RSGTETDTIADIALVNCGOIKTCAPARSDRVAKYNQLLRIEQLLDGAGVYAGRSAPPR 422
 DB 364 RSGTETDTMIADLVAIGSGOIKTCAPASRSRVAKYNQLLRIEALGDGAARYAGDLAPPR 423

QY 423 F 423
 DB 424 F 424

RESULT 10
 Q6AAB8 PRELIMINARY; PRT; 426 AA.

ID Q6AAB8
 AC Q6AAB8;
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
 DE Enolase (EC 4.2.1.11).
 GN OrderedLocusNames=PFA0545;
 OS Propionibacterium acnes.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Propionibacteriaceae; Propionibacteriaceae; Propionibacterium.

AC Q9CDA2; 2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 45, Last annotation update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE Enolase (EC 4.2.1.11) (2-phosphoglycerate dehydratase) (2-phospho-D-
 DE glycerate hydro-lyase).
 GN Name: eno; OrderedLocusNames=ML0255;
 GN Name: eno; OrderedLocusNames=ML0255;
 OS Mycobacterium lepreae.
 OS Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1769;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=TN;
 RX MEDLINE=21128732; PubMed=11234002; DOI=10.1038/35059006;
 RA Cole S.T., Biglmeier K., Parkhill J., James K.D., Thomson N.R.,
 RA Wheeler P.R., Honore N., Garnier T., Chillingworth T., Connor R.,
 RA Mungall K.L., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R.M., Devlin K., Duthoy S., Feltwell T., Fraser A., Hamlin N.,
 RA Holroyd S., Hornsby T., Jagels K., Lacroix C., Maclean J., Moule S.,
 RA Murphy L.D., Oliver K., Quail M.A., Rajadream M.A., Rutherford K.M.,
 RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
 RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
 RA Barrell B.G.;
 RT "Massive gene decay in the leprosy bacillus.";
 RL Nature 409:1007-1011(2001).
 CC -1- CATALYTIC ACTIVITY: 2-phospho-D-glycerate = phosphoenolpyruvate +
 CC H(2O).
 CC -1- COFACTOR: Magnesium is required for catalysis and for stabilizing
 CC the dimer (By similarity).
 CC -1- PATHWAY: Glycolysis.
 CC -1- SUBUNIT: Homodimer (By similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: Belongs to the enolase family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announcements/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AL593917; CAC29763.1; ALT_INIT.
 CC HSPF; Q9NDH8; 10EP.
 CC Leproma; ML0255; -; 1.
 CC HAMAP; MF_00318; -; 1.
 CC InterPro; IPR000941; Enolase.
 CC Pfam; PF00113; Enolase_C; 1.
 CC Pfam; PF03952; Enolase_N; 1.
 CC PRINTS; PR00148; ENOLASE.
 CC ProDom; PD000902; Enolase; 1.
 CC TIGRFAMs; TIGR01060; eno; 1.
 CC PROSITE; PS00164; ENOLASE; 1.
 CC Complete proteome; Glycolysis; Lyase; Magnesium.
 CC ACT_SITE 154 154 By similarity.
 CC METAL 241 241 Magnesium (By similarity).
 CC METAL 283 283 Magnesium (By similarity).
 CC METAL 310 310 Magnesium (By similarity).
 CC SEQUENCE 429 AA; 45201 MW; A8299C1EA44617B5 CRC64;
 CC Query Match 69.6%; Score 1500.5; DB 1; Length 429;
 CC Best Local Similarity 71.0%; Pred No. 1.2e-80;
 CC Matches 299; Conservative 47; Mismatches 74; Indels 1; Gaps 1;
 QY 4 IMHVFAREILDSRGNTPTVEAEVFLDDGSHGVAGVPSGASTGVHEAHLRDGDRYLKGV 63
 Db 4 IEQVGAREILDSRGNTPTVEAEVFLDDGSHGVAGVPSGASTGVHEAHLRDGDRYLKGV 63
 QY 64 LKAVENVEETGDELAGEADQRLIDQELDLDTGTPDKSRIGGNAILGVSAVAKAAAD 123
 Db 64 KRAVDVILDEIGPVVIGLGNANDQRLIDQELDLDTGTPDKSRIGGNAILGVSAVAKAAAD 123

QY 124 SAGLPLFRYIGGNPAHLVPPMMNIINGAHADSGVDQEFMIPIGAEFTFSEALNGAE 183
 Db 124 SHELPLFRYIGGSNAHLVPPMMNITLNGAHADTAVDQEFNVAPIGAPSFVEALRWGAE 183
 QY 184 VYHALKSVIKKGLSTGLDEGGPAPSGVSTREALDLVLEALEKAGFTTPGKDIALALDVA 243
 Db 184 VYHALKSVLKKGSLSTGLDEGGPAPSGVSTREALDLVLEALEKAGFTTPGKDIALALDVA 243
 QY 244 SSEFFKDGIT-VHPEGGQSHAAEMANVYAEVLVDAYPIVSTIEDPLQDDWEGYTNLTATIGD 302
 Db 244 ATEFYTDGIGYHFGHTADQMTFADLLGSYPLVSLIEDPLSEDDWDGWAALTASIGE 303
 QY 303 KQIVGDDDFVTNPERLKEGIAKKAANSILVKVNIQTGLTETFDVDMAHRAGYTSMMSH 362
 Db 304 QVIVGDDDFATNPERLESGIGRVANALLVKVNIQTGLTETLEAVALAHSGYRTMISH 363
 QY 363 RSGETEDTTIADLVALNCGQIKTGAPARSDRVAKYNQLRLIEQLLDGAGVAGSAPFR 422
 Db 364 RSGETEDTTIADLVALNCGQIKTGAPARSDRVAKYNQLRLIEQLLDGAGVAGSAPFR 422
 QY 423 F 423
 Db 424 Y 424
 RESULT 12
 Q83H73 PRELIMINARY; PRT; 428 AA.
 AC Q83H73;
 DT 01-JUN-2003 (TREMBlrel. 24, Created)
 DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
 DE Enolase (EC 4.2.1.11).
 GN Name: eno; OrderedLocusNames=TW93;
 OS Tropheryma whipplei (strain TW08/27) (Whipple's bacillus).
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Micrococciaceae; Cellulomonadaceae; Tropheryma.
 OX NCBI_TaxID=218496;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=TW08/27;
 RX MEDLINE=22495039; PubMed=12606174; DOI=10.1016/S0140-6736(03)12597-4;
 RA Bentley S.D., Mairwald M., Murphy L.D., Pallen M.J., Yeats C.A.,
 RA Dover L.G., Norbertczak H.T., Besra G.S., Quail M.A., Harris D.E.,
 RA von Herbay A., Goble A., Rutter S., Squares R., Squares S.,
 RA Barrell B.G., Parkhill J., Rellman D.A.;
 RT "Sequencing and analysis of the genome of the Whipple's disease
 RT bacterium Tropheryma whipplei.";
 RL Lancet 361:637-644(2003).
 CC -1- CATALYTIC ACTIVITY: 2-phospho-D-glycerate = phosphoenolpyruvate +
 CC H(2O).
 CC -1- COFACTOR: Magnesium is required for catalysis and for stabilizing
 CC the dimer (By similarity).
 CC -1- PATHWAY: Glycolysis.
 CC -1- SUBUNIT: Homodimer (By similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: Belongs to the enolase family.
 CC EMBL; BX251412; CAB67452.1; -.
 CC HSPF; Q9NDH8; 10EP.
 CC GO; GO:0000015; C:phosphopyruvate hydratase complex; IEA.
 CC GO; GO:0016829; F:lyase activity; IEA.
 CC GO; GO:0000287; F:magnesium ion binding; IEA.
 CC GO; GO:0004634; F:phosphopyruvate hydratase activity; IEA.
 CC GO; GO:0006096; P:glycolysis; IEA.
 CC InterPro; IPR000941; Enolase.
 CC Pfam; PF00113; Enolase_C; 1.
 CC Pfam; PF03952; Enolase_N; 1.
 CC PRINTS; PR00148; ENOLASE.
 CC ProDom; PD000902; Enolase; 1.
 CC TIGRFAMs; TIGR01060; eno; 1.
 CC PROSITE; PS00164; ENOLASE; 1.
 CC Complete proteome; Glycolysis; Lyase; Magnesium.
 CC SEQUENCE 428 AA; 45961 MW; 356D63DCCC24B581 CRC64;
 SQ

```

Query Match      66.6%; Score 1436; DB 2; Length 428;
Best Local Similarity 67.9%; Pred. No. 7.7e-77;
Matches 286; Conservative 51; Mismatches 82; Indels 2; Gaps 2;

Qy 7 VFAREILDSRGNPTVEAEVFLDDSGHVAGVPSGASTGVHAEHLRD-GGDRYLKGVK 65
Db 7 VHAQIILDSRGTPTEVEVTLDEDSMGRSVPSGASTGAFAEHLRDQDNNEYLKGVTR 66
Qy 66 AVENVNEEIGDELAGLEADDQRLIDEAMI KLDGTANKSRIGCANAILGVSMVAKAAADS 125
Db 67 AVRSVNSEIAPVLIGFADFQVGLDHRIEIDLTGNNKSRIGCANAILGVSLASAAARA 126
Qy 126 GLPLFRYIGGNPAHVLFPVPMNIIINGGAHDSGVDOEFMIAPICAEFTFSEALRNGAEVY 185
Db 127 DLSLFRYLGGPSSRILFPVPMNIIINGGAHADTGVDIQEFMILPVGARSFSESRLWGCEVY 186
Qy 186 HALKSIVKEKLSGLGDEGGFAPSVGSTREALDLIVEAIKAGFTPGKDTALALDVASS 245
Db 187 HSLKVQLRESGLSGDEGGFAPALRSNRRTALDLIILSAIEKAGFSPGIDIVLALDIAAS 246
Qy 246 EFKF-DGTYHFEFGQHSAAEMANVYAEIADYAPIVSTEDPLQEDDWEQYTNLTATIGDKV 304
Db 247 EFKCAPGHYRFEKGDIITSDELISYEGLLSYPLVSIEDPLQDDWEGYRTLTTHIGDRV 306
Qy 305 QIVGDDFFVTNPERLKEGIAKKAANSILVKVQNIQIGTLTETFDVADMAHRAGYTTSMMSHRS 364
Db 307 QIVGDDLVFVNTSRLSRGIQSGVANSILVKVQNIQIGTLTETFDVADMAKGGYTAVALSHRS 366
Qy 365 GETEDTTIADLAVALNCQIKTGAPARSDRVAKYNQLLRIBQLLGDAGVYAGRSAPFRFQ 424
Db 367 GETEDTTIADMAVATNCGQIKTGAPARGERTAKYNQLLRIBEKLGRSARYAGWLSYPRWQ 426
Qy 425 G 425
Db 427 G 427

RESULT 13
Q83FF7 PRELIMINARY; PRT; 446 AA.
AC Q83FF7;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Enolase (EC 4.2.1.11).
GN Name=eno; OrderedLocusNames=TW783;
OS Tropheryma whipplei (strain Twist) (Whipple's bacillus).
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Micrococciaceae; Cellulomonadaceae; Tropheryma.
OX NCBI_TaxID=203267;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Twist;
RA MEDLINE=22784088; PubMed=12902375;
RA Raoult D., Ogata H., Audic S., Robert C., Suhre K., Drancourt M.,
RA Claverie J.-M.;
RT "Tropheryma whipplei Twist: a human pathogenic Actinobacteria with a
RT reduced genome.";
RL Genome Res. 13:1800-1809 (2003).
CC -|- CATALYTIC ACTIVITY: 2-phospho-D-glycerate = phosphoenolpyruvate +
CC H(2)O.
CC -|- COFACTOR: Magnesium is required for catalysis and for stabilizing
CC the dimer (By similarity).
CC -|- PATHWAY: Glycolysis.
CC -|- SUBUNIT: Homodimer (By similarity).
CC -|- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -|- SIMILARITY: Belongs to the enolase family.
DR EMBL: AE016852; AA04880.1; -.
DR HSP: Q9NDH8; 10EP.
DR GO: GO:000015; C:phosphopyruvate hydratase complex; IEA.
DR GO: GO:001629; F:lyase activity; IEA.
DR GO: GO:0000287; F:magnesium ion binding; IEA.
DR GO: GO:0004634; F:phosphopyruvate hydratase activity; IEA.
DR GO: GO:0006096; P:glycolysis; IEA.
DR InterPro: IPR000941; Enolase.
DR Pfam: PF00113; Enolase_C; 1.
DR Pfam: PF03952; Enolase_N; 1.
DR PRINTS: PR00148; ENOLASE.
DR PRODOM: PD000902; Enolase; 1.
DR TIGRFAMs: TIGR01060; eno; 1.
DR PROSITE: PS00164; ENOLASE; 1.
KW Complete proteome; Glycolysis; Lyase; Magnesium.
SQ SEQUENCE 446 AA; 48141 MW; 786E73F99F94A6DA CRC64;

Query Match      66.6%; Score 1436; DB 2; Length 446;
Best Local Similarity 67.9%; Pred. No. 8e-77;
Matches 286; Conservative 51; Mismatches 82; Indels 2; Gaps 2;

Qy 7 VFAREILDSRGNPTVEAEVFLDDSGHVAGVPSGASTGVHAEHLRD-GGDRYLKGVK 65
Db 25 VHAQIILDSRGTPTEVEVTLDEDSMGRSVPSGASTGAFAEHLRDQDNNEYLKGVTR 84
Qy 66 AVENVNEEIGDELAGLEADDQRLIDEAMI KLDGTANKSRIGCANAILGVSMVAKAAADS 125
Db 85 AVRSVNSEIAPVLIGFADFQVGLDHRIEIDLTGNNKSRIGCANAILGVSLASAAARA 144
Qy 126 GLPLFRYIGGNPAHVLFPVPMNIIINGGAHDSGVDOEFMIAPICAEFTFSEALRNGAEVY 185
Db 145 DLSLFRYLGGPSSRILFPVPMNIIINGGAHADTGVDIQEFMILPVGARSFSESRLWGCEVY 204
Qy 186 HALKSIVKEKLSGLGDEGGFAPSVGSTREALDLIVEAIKAGFTPGKDTALALDVASS 245
Db 205 HSLKVQLRESGLSGDEGGFAPALRSNRRTALDLIILSAIEKAGFSPGIDIVLALDIAAS 264
Qy 246 EFKF-DGTYHFEFGQHSAAEMANVYAEIADYAPIVSTEDPLQEDDWEQYTNLTATIGDKV 304
Db 265 EFKCAPGHYRFEKGDIITSDELISYEGLLSYPLVSIEDPLQDDWEGYRTLTTHIGDRV 324
Qy 305 QIVGDDFFVTNPERLKEGIAKKAANSILVKVQNIQIGTLTETFDVADMAHRAGYTTSMMSHRS 364
Db 325 QIVGDDLVFVNTSRLSRGIQSGVANSILVKVQNIQIGTLTETFDVADMAKGGYTAVALSHRS 384
Qy 365 GETEDTTIADLAVALNCQIKTGAPARSDRVAKYNQLLRIBQLLGDAGVYAGRSAPFRFQ 424
Db 385 GETEDTTIADMAVATNCGQIKTGAPARGERTAKYNQLLRIBEKLGRSARYAGWLSYPRWQ 444
Qy 425 G 425
Db 445 G 445

RESULT 14
ENO_AGRT5
ID -ENO_AGRT5 STANDARD; PRT; 424 AA.
AC Q8UFH1;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Enolase (EC 4.2.1.11) (2-phosphoglycerate dehydratase) (2-phospho-D-
DE glycerate hydro-lyase).
GN Name=eno; OrderedLocusNames=Atu1426, AGR_C.2631;
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
OX NCBI_TaxID=176299;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608550; PubMed=11743193; DOI=10.1126/science.1066804;
RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. Sr.,
RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
RA Kutyavin T., Levy R., Li M.-J., McClelland E., Palmeri A.,
RA Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D.,
RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,

```

	QY	244	SSEFFKDGTYHFEGGSHS--AAEMANVYAELVDAYPIVSIEDPLOBDDWEGYTNLTAIG	301
	Dd	244	STEFFPDGGKYVLGEGRITLPCAMAEYLAELVNKYPILISVEDGMADDWEGWKTLIDLAVG	303
	QY	302	GKVIVGDGDFVTNPNRLKEGIKKKAANSILVKVNQIGLTFTTFDAVMAHRAGYTSMMS	361
	Dd	304	NKCQLVGDDLFTNSARLRDGIKMGVANSILVKVNQIGLSSETLDVAETAHKAGYTVAMS	363
	QY	362	HRSGETEDTTIADLAVALNCGOIKTGAPARSDRVAKYNOLLRIEQLLGDAGVVYAGRS	418
	Dd	364	HRSGETEDTTIADLAVALNCGOIKTGSUARSRLRKLAKYNOLIRIEEMLGPQAAYAGRS	420
 RESULT_15 ENO_THETN STANDARD; PRT; 429 AA.				
	AC	QR967;		
	DT	28-FEB-2003 (Rel. 41, Created)		
	DR	28-FEB-2003 (Rel. 41, Last sequence update)		
	DT	25-OCT-2004 (Rel. 45, Last annotation update)		
	DE	Enolase (EC 4.2.1.11) (2-phosphoglycerate dehydratase) (2-phospho-D-		
	DE	glycerate hydro-lyase).		
	GN	Name=eno; OrderedLocusNames=TTEL1759;		
	OS	Thermoanaerobacter tengcongensis.		
	OC	Bacteria; Firmicutes; Clostridia; Thermoanaerobacterales;		
	OX	Thermoanaerobacteriaceae; Thermoanaerobacter.		
	QC	NCBI_TaxID=119072;		
	RN	[1]		
	RP	SEQUENCE FROM N.A.		
	RX	STRAIN=MBA / JCM 11007;		
	RX	MEDLINE=2192816; PubMed=11997336; DOI=10.1101/gr.219302;		
	RA	Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,		
	RA	Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,		
	RA	Tan H., Chen R., Wang J., Yu J., Yang H.;		
	RA	"A complete sequence of the T. tengcongensis genome.";		
	RL	Genome Res. 12:1689-700(2002).		
	CC	-[- CATALYTIC ACTIVITY: 2-phospho-D-glycerate = phosphoenolpyruvate +		
	CC	H(2O).]		
	CC	-[- COFACTOR: Magnesium is required for catalysis and for stabilizing		
	CC	the dimer (By similarity).]		
	CC	-[- PATHWAY: Glycolysis.]		
	CC	-[- SUBUNIT: Homodimer (By similarity)]		
	CC	-[- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).]		
	CC	-[- SIMILARITY: Belongs to the enolase family.]		
	CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
	CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation at		
	CC	the European Bioinformatics Institute. There are no restrictions on its		
	CC	use by non-profit institutions as long as its content is in no way		
	CC	modified and this statement is not removed. Usage by and for commercial		
	CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/		
	CC	or send an email to licensee@isb-sib.ch).		
	CC	-----		
	CC	ENBL; AE013129; AAM24953.1; -.		
	DR	HSSP; P08324; IE91.		
	DR	HMAP; MF 00318; -. 1.		
	DR	InterPro; IPR000941; Enolase.		
	DR	Pfam; PF00113; Enolase_C; 1.		
	DR	Pfam; PF03952; Enolase_N; 1.		
	DR	PRINTS; PR00148; ENOLASE.		
	DR	ProDom; PD000902; Enolase; 1.		
	DR	TIGRFAMs; TIGR01060; eno; 1.		
	DR	PROSITE; PS00164; ENOLASE; 1.		
	KW	Complete proteome; Glycolysis; Lyase; Magnesium.		
	FT	ACT SITE 155 155 By similarity.		
	FT	METAL 242 242 Magnesium (By similarity).		
	FT	METAL 286 286 Magnesium (By similarity).		
	FT	METAL 313 313 Magnesium (By similarity).		
	FT	METAL 313 313 Magnesium (By similarity).		
	SC	SEQUENCE 429 AA; 46296 MW; ABCC244SDBEA8D65 CRC64;		
 Query Match Similarity 65.5%; Score 1412; DB 1; Length 429; Best Local Similarity 66.3%; Pred. No. 2e-75;				

Matches		281;	Conservative	57;	Mismatches	82;	Indels	4;	Gaps	3;
Qy	1	VAEIMHVFAREILDSRGNPTVEAEVFLDDGSHGVAGVPSGASTGVHEAHELDDG-DRYL 59								
Db	1	MSSIIDIYAREILDSRGNPTIEVEVELDSGAVGRAAVPSGASTGAFAEIELRDGDKSRYL 60								
Qy	60	GKGVLKAVENNEEIGDELAGLEADDQRLIDEAMIKLDGTANKSRGLGANAILGVSMVAK 119								
Db	61	GKGVLKAVQNVNDIIAPELIGMEAQDQVAIDKAMIELDGTENKSKLGANAILGVSLAVAK 120								
Qy	120	AAADSAGLPLFRYIGGNPAHVLVPMNNIINGGAHADSGVDVQEFMPIAPIGAETFSALR 179								
Db	121	AAAEELGLPLYQYLGGVNACTLPVPMNNILNGGKHADNNVDIQEFMIMPVGPAPNFKEARL 180								
Qy	180	NGAEVYHALKSVIKEKGLSTGLGDEGGFAPSVGSTREALDLIVEAIEKAGFTPGKDIALA 239								
Db	181	MCSEVYHSLKNVLHSLKGLSTVGVDEGGFAPNLTSNEEAIKVILEAIEKAGVYVPGEDIVLA 240								
Qy	240	LDVASSBFFK-DGTYHFEKG--QHSAAEAMNVAELVDAYPIVSIEDPQEDDWEGYTNL 296								
Db	241	LDPAAATEMYKEDGKYHFEKGIVRTSEEMIEFWEQLVSKYPIVSIEDGLAEEDWNGKLL 300								
Qy	297	TATIGDKVQIVGDDFFVTNPERLKEGIAKKAANSILVKVNOIGTLTETTFDAVDMAHKAGY 356								
Db	301	TERLGKKVQLVGDDLFVTNTERLSKGINMGVANSILIKLNQIGTLTETDLDAIENAKRAGY 360								
Qy	357	TSMMSHRSGETEDTTIADLAVALNCGQIKTGAPARSDRVAKYNOLLRIEQLLGDAGVYAG 416								
Db	361	TAIVSHRSGETEDTTIADLVVATNVGQIKTGAPARTDRVAKYNOLLRIEALGSAVOYPG 420								
Qy	417	RSAF 420								
Db	421	KNAF 424								

Search completed: September 29, 2005, 12:13:46
Job time : 92.5 secs

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 28, 2005, 20:45:32 ; Search time 4241.5 Seconds
(without alignments)
2568.796 Million cell updates/sec

Title: US-10-728-947-1
Perfect score: 1578
Sequence: 1 ggctgggagatgggtagtt.....ctcaagcagggaactgctt 1578

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 7442561 seqs, 3452328358 residues

Total number of hits satisfying chosen parameters: 14885122

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US05_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:*
- 18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq:*
- 19: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq:*
- 20: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq:*
- 21: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq:*
- 22: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 23: /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq:*
- 24: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
- 25: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 26: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1578	100.0	1578	9	US-09-860-768-1
2	1578	100.0	1578	20	US-10-728-947-1
3	1578	100.0	3309400	9	US-09-738-626-1
4	1576.4	99.9	1578	20	US-09-860-768-3
5	1576.4	99.9	1578	20	US-10-728-947-3
6	1403.4	88.9	1405	21	US-10-494-836-7
7	1396.4	88.5	1398	19	US-10-781-014-71

8	1275	80.8	1275	9	US-09-738-626-1085
9	908.6	57.6	1275	17	US-10-282-122A-17724
10	698.4	44.3	9025608	15	US-10-156-761-1
11	697	44.2	1284	15	US-10-156-761-3520
12	679.8	43.1	1269	17	US-10-369-493-32027
13	647	41.0	1287	17	US-10-282-122A-25949
14	632.6	40.1	45191	15	US-10-080-170-649
15	632.6	40.1	45191	19	US-10-468-356-649
16	632.6	40.1	45191	19	US-10-468-356-649
17	632.4	40.1	1317	17	US-10-282-122A-26512
18	626.2	39.7	1290	17	US-10-282-122A-28306
19	584.6	37.0	2256646	19	US-10-470-565-1
20	583.4	37.0	1344	17	US-10-282-122A-27498
21	582	36.9	1269	17	US-10-369-493-34244
22	575.4	36.5	1269	17	US-10-369-493-35383
23	574	36.4	1275	17	US-10-369-493-35506
24	574	36.4	1281	17	US-10-369-493-40598
25	569.8	36.1	1278	17	US-10-369-493-41573
26	569	36.1	1272	17	US-10-369-493-43439
27	567.4	36.0	1260	17	US-10-369-493-38436
28	567.4	36.0	1269	17	US-10-369-493-38909
29	561.8	35.6	1272	17	US-10-369-493-31449
30	557	35.3	1278	17	US-10-369-493-45417
31	552	35.0	1290	17	US-10-369-493-43178
32	542	34.3	1281	17	US-10-282-122A-11620
33	531.8	33.7	1290	9	US-09-815-242-7834
34	531.8	33.7	1290	17	US-10-282-122A-30326
35	530.8	33.6	1281	17	US-10-369-493-44693
36	528	33.5	1287	17	US-10-369-493-32181
37	522.8	33.1	1278	17	US-10-369-493-28375
38	522.8	33.1	1281	17	US-10-369-493-31134
39	522.8	33.1	1281	17	US-10-282-122A-13098
40	515.2	32.6	1404	17	US-10-282-122A-33598
41	514.8	32.6	1269	17	US-10-369-493-47120
42	513.6	32.5	1293	17	US-10-369-493-46986
43	509.4	32.3	1293	9	US-10-369-493-37536
44	507.6	32.2	1293	9	US-09-974-300-664
45	504.8	32.0	1293	17	US-10-282-122A-9983

ALIGNMENTS

RESULT 1
US-09-860-768-1
; Sequence 1, Application US/09860768
; Patent No. US20020082403A1
; GENERAL INFORMATION:
; APPLICANT: Mockel, Bettina
; APPLICANT: Pfeifferle, Walter
; APPLICANT: Hermann, Thomas
; APPLICANT: Pohler, Alfred
; APPLICANT: Kalinowski, Jörn
; APPLICANT: Bathe, Brigitte
; TITLE OF INVENTION: New Nucleotide Sequences that Code for the Eno Gene
; FILE REFERENCE: 21123/278404
; CURRENT APPLICATION NUMBER: US/09/860.768
; CURRENT FILING DATE: 2001-05-21
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 1578
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
US-09-860-768-1

Query Match 100.0%; Score 1578; DB 9; Length 1578;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1578; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GGCTGGGATGGGTAGTTTCGCCACTAATTTCACTGATTCCTCATCGAACAGA 60
Db 1 GGCTGGGATGGGTAGTTTCGCCACTAATTTCACTGATTCCTCATCGAACAGA 60

QY 61 TTCTGTCACAAATTCGGTGTAGACGTGATTTGAAGACATTTGATCAGGTGAATAATTTCTAG 120
 Db 61 TTCTGTCACAAATTCGGTGTAGACGTGATTTGAAGACATTTGATCAGGTGAATAATTTCTAG 120
 QY 121 TTAGCTCCCAAGTTGGCTAGAGGCGCCACAGTGGCTGAAATCATGACAGTATTCGCTCGC 180
 Db 121 TTAGCTCCCAAGTTGGCTAGAGGCGCCACAGTGGCTGAAATCATGACAGTATTCGCTCGC 180
 QY 181 GAAATTCGACATCCCGGCTTAAACCCAAACCGTTCGAGGCGAGAGTTTCTGATGACGGT 240
 Db 181 GAAATTCGACATCCCGGCTTAAACCCAAACCGTTCGAGGCGAGAGTTTCTGATGACGGT 240
 QY 241 TCCCAAGGTTGCGAGGTTTCCATCCGCGCATCCACGGGCTCCACGAGGCTCATGAG 300
 Db 241 TCCCAAGGTTGCGAGGTTTCCATCCGCGCATCCACGGGCTCCACGAGGCTCATGAG 300
 QY 301 CTGCGTGAAGGTTGCGATCGCTTACCTGGGCAAGGCGTTTGAAGGCGAGTTGMAAACGTC 360
 Db 301 CTGCGTGAAGGTTGCGATCGCTTACCTGGGCAAGGCGTTTGAAGGCGAGTTGMAAACGTC 360
 QY 361 AACGAAGAAATCGGCGATCGCTTACCTGGGCAAGGCGTTTGAAGGCGAGTTGMAAACGTC 420
 Db 361 AACGAAGAAATCGGCGATCGCTTACCTGGGCAAGGCGTTTGAAGGCGAGTTGMAAACGTC 420
 QY 421 GAAGCAATGATCAAGCTTTGATGCGACCGCGCAACCAAGTCCCGCTGGGTGCAACGCAATC 480
 Db 421 GAAGCAATGATCAAGCTTTGATGCGACCGCGCAACCAAGTCCCGCTGGGTGCAACGCAATC 480
 QY 481 CTGCGTGAAGGTTGCGATCGCTTACCTGGGCAAGGCGTTTGAAGGCGAGTTGMAAACGTC 540
 Db 481 CTGCGTGAAGGTTGCGATCGCTTACCTGGGCAAGGCGTTTGAAGGCGAGTTGMAAACGTC 540
 QY 541 CGCTACATCGGTGGAACCAACGACACAGTCTTCTCCAGTTTCCAAATGATGAACATCATCAAC 600
 Db 541 CGCTACATCGGTGGAACCAACGACACAGTCTTCTCCAGTTTCCAAATGATGAACATCATCAAC 600
 QY 601 GGTGCGCTCAGCTGATCTCGGCTGTGACGTTGAGGTTGAGGATTCATGATCGCTCCAAATCGT 660
 Db 601 GGTGCGCTCAGCTGATCTCGGCTGTGACGTTGAGGTTGAGGATTCATGATCGCTCCAAATCGT 660
 QY 661 GCAGAGACCTTCTGAGGCTCTCGCAACGCGCGGAGGTTTACACGCACTGAAGTCC 720
 Db 661 GCAGAGACCTTCTGAGGCTCTCGCAACGCGCGGAGGTTTACACGCACTGAAGTCC 720
 QY 721 GTTCATCAAGAAAGGCGCTTCCACCGGACTTGGCGATGAGGCGGCTTGGCTCTTCC 780
 Db 721 GTTCATCAAGAAAGGCGCTTCCACCGGACTTGGCGATGAGGCGGCTTGGCTCTTCC 780
 QY 781 GTGCGCTCACCCTGAGGCTCTTGAACCTTATCGTTGAGGCAATCGAGAGGCTGGCTTC 840
 Db 781 GTGCGCTCACCCTGAGGCTCTTGAACCTTATCGTTGAGGCAATCGAGAGGCTGGCTTC 840
 QY 841 ACCCAGGCAAGGACATCGCTCTGCTGAGGCTTCTCTCTGAGTTCTTCAAGGAC 900
 Db 841 ACCCAGGCAAGGACATCGCTCTGCTGAGGCTTCTCTCTGAGTTCTTCAAGGAC 900
 QY 901 GGCACCTTACCACTTCCAGAGTGGCGGACGACTCCGAGCTGAGTGGCAACCGTTTACGCT 960
 Db 901 GGCACCTTACCACTTCCAGAGTGGCGGACGACTCCGAGCTGAGTGGCAACCGTTTACGCT 960
 QY 961 GAGCTGTTGACGCTGATCCCAATCGTCTCATCGAGGACCCGCTGAGGAGATGACTGG 1020
 Db 961 GAGCTGTTGACGCTGATCCCAATCGTCTCATCGAGGACCCGCTGAGGAGATGACTGG 1020
 QY 1021 GAGGTTACACCAACTCAGGCTGAGGAGGCGATCGCTTAAGAGGCTGCCAACTCC 1080
 Db 1021 GAGGTTACACCAACTCAGGCTGAGGAGGCGATCGCTTAAGAGGCTGCCAACTCC 1080
 QY 1081 TTCTTCGTCAACCAACTCAGGCTGAGGAGGCGATCGCTTAAGAGGCTGCCAACTCC 1140
 Db 1081 TTCTTCGTCAACCAACTCAGGCTGAGGAGGCGATCGCTTAAGAGGCTGCCAACTCC 1140

RESULT 2

US-10-728-947-1
 ; Sequence 1 Application US/10728947
 ; Publication No. US20040220394A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Mockel, Bettina
 ; APPLICANT: Pfeifferle, Walter
 ; APPLICANT: Hermann, Thomas
 ; APPLICANT: Pohler, Alfred
 ; APPLICANT: Kalinowski, Jorn
 ; APPLICANT: Bache, Brigitte
 ; TITLE OF INVENTION: New Nucleotide Sequences that Code for the Eno Gene
 ; FILE REFERENCE: 21123/278404
 ; CURRENT APPLICATION NUMBER: US/10728,947
 ; CURRENT FILING DATE: 2003-12-08
 ; PRIOR APPLICATION NUMBER: US/09/860,768
 ; PRIOR FILING DATE: 2001-05-21
 ; NUMBER OF SEQ ID NOS: 6
 ; SOFTWARE: Patent In version 3.0
 ; SEQ ID NO 1
 ; LENGTH: 1578
 ; TYPE: DNA
 ; ORGANISM: Corynebacterium glutamicum
 ; US-10-728-947-1
 Query Match 100.0%; Score 1578; DB 20; Length 1578;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1578; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGCTGGGGAATAGGGTATGTTTCCGCACTAATTTCAACTGATTCGCTCATCGAAACAAGA 60
 Db 1 GGCTGGGGAATAGGGTATGTTTCCGCACTAATTTCAACTGATTCGCTCATCGAAACAAGA 60
 QY 61 TTCTGTCACAAATTCGGTGTAGACGTGATTTGAAGACATTTGATCAGGTGAATAATTTCTAG 120
 Db 61 TTCTGTCACAAATTCGGTGTAGACGTGATTTGAAGACATTTGATCAGGTGAATAATTTCTAG 120
 QY 121 TTAGCTCCCAAGTTGGCTAGAGGCGCCACAGTGGCTGAAATCATGACAGTATTCGCTCGC 180
 Db 121 TTAGCTCCCAAGTTGGCTAGAGGCGCCACAGTGGCTGAAATCATGACAGTATTCGCTCGC 180

181 GAAATTTCTGACTCCCGCGGTACCAACCGTTCGAGCGAGAGGTTTCTTGGATGACGGT 240
 181 GAAATTTCTGACTCCCGCGGTACCAACCGTTCGAGCGAGAGGTTTCTTGGATGACGGT 240
 241 TCCACCGGTTCGACGGTTCCTATCCGCGCATCCACCGCGGTTCACGAGGCTCATGAG 300
 241 TCCACCGGTTCGACGGTTCCTATCCGCGCATCCACCGCGGTTCACGAGGCTCATGAG 300
 301 CTGCGTGAAGGTGGCCATCGCTACCTTGGCGAAGGGGCTTTTGAAGCAGTTGAAGCGTC 360
 301 CTGCGTGAAGGTGGCCATCGCTACCTTGGCGAAGGGGCTTTTGAAGCAGTTGAAGCGTC 360
 361 AACGAAGAAATCGGCGACGAGCTCGCTTGGCTTAGAGGCTGACGATCAGCGCTTCATCGAC 420
 361 AACGAAGAAATCGGCGACGAGCTCGCTTGGCTTAGAGGCTGACGATCAGCGCTTCATCGAC 420
 421 GAAGCAATGATCAAGCTTGAATGSCACCGCCAAACAGTCCCGCTCGGTCGAAACGCAATC 480
 421 GAAGCAATGATCAAGCTTGAATGSCACCGCCAAACAGTCCCGCTCGGTCGAAACGCAATC 480
 481 CTTGGTGTTCATGCTGTGCAAAAGGCTGCTGCTGATTCGCGAGGCTCCCACTGTC 540
 481 CTTGGTGTTCATGCTGTGCAAAAGGCTGCTGCTGATTCGCGAGGCTCCCACTGTC 540
 541 CGCTACATCGGTGGACCAACGACACAGCTTCTTCCAGTTCCTCAATGATGAACATCATCAAC 600
 541 CGCTACATCGGTGGACCAACGACACAGCTTCTTCCAGTTCCTCAATGATGAACATCATCAAC 600
 601 GGTGGCGCTCACGCTGACTCCGCTGTTGAGCTTCAGGAAATTCATGATCGCTCCCAATCGGT 660
 601 GGTGGCGCTCACGCTGACTCCGCTGTTGAGCTTCAGGAAATTCATGATCGCTCCCAATCGGT 660
 661 GCAGAGACCTTCTCTGAGGCTTCGCGAAGCGGCGGAGGCTTACCAAGCTGAGTCC 720
 661 GCAGAGACCTTCTCTGAGGCTTCGCGAAGCGGCGGAGGCTTACCAAGCTGAGTCC 720
 721 GTCATCAAGAAAGGCGCTTCCACCGGCTTGGCGATGAGGCGGCTTCCTCTCC 780
 721 GTCATCAAGAAAGGCGCTTCCACCGGCTTGGCGATGAGGCGGCTTCCTCTCC 780
 781 GTCCGCTCACCGCTGAGGCTCTTGAACCTTATCGTTGAGGCAATCAGAGGCTGCTTC 840
 781 GTCCGCTCACCGCTGAGGCTCTTGAACCTTATCGTTGAGGCAATCAGAGGCTGCTTC 840
 841 ACCCAGGCAAGACATCGCTTCTGAGCTTCGAGCTTCTCTGAGTTCCTTCAAGGAC 900
 841 ACCCAGGCAAGACATCGCTTCTGAGCTTCGAGCTTCTCTGAGTTCCTTCAAGGAC 900
 901 GGCACCTACCACTTCGAAGGTGGCGAGCACTCCGAGCTGAGATGCGCAACGTTTACGCT 960
 901 GGCACCTACCACTTCGAAGGTGGCGAGCACTCCGAGCTGAGATGCGCAACGTTTACGCT 960
 961 GAGCTGTTGACGCGTACCAATCGCTCTCCATCGAGGACCACTGCGAGGAGATGACTGG 1020
 961 GAGCTGTTGACGCGTACCAATCGCTCTCCATCGAGGACCACTGCGAGGAGATGACTGG 1020
 1021 GAGGTTTACCACTTCACCGCAACCATTCGCGGCAAGGTTTCAAGGCTGCGCAACGAC 1080
 1021 GAGGTTTACCACTTCACCGCAACCATTCGCGGCAAGGTTTCAAGGCTGCGCAACGAC 1080
 1081 TTCTTCGTACCAACCTGAGCGCTTGAAGGAGGCGATCGCTTGAAGGCTGCGCAACGAC 1140
 1081 TTCTTCGTACCAACCTGAGCGCTTGAAGGAGGCGATCGCTTGAAGGCTGCGCAACGAC 1140
 1141 ATCTCTGTTAAGTGAACAGATCGCTTACCTCAGGAGGCTTGAAGGCTGCGCAACGAC 1200
 1141 ATCTCTGTTAAGTGAACAGATCGCTTACCTCAGGAGGCTTGAAGGCTGCGCAACGAC 1200
 1201 GCTTACCGGCGAGGCTTACCTTCCATGATGTCCTCAGGCTTCCGCTGAGACCGGAGCAC 1260
 1201 GCTTACCGGCGAGGCTTACCTTCCATGATGTCCTCAGGCTTCCGCTGAGACCGGAGCAC 1260

1261 ACCATTGCTGACCTCGAGTTGCACTCAACTGTGGCCAGATCAAGACTGGTCTCCAGCA 1320
 1261 ACCATTGCTGACCTCGAGTTGCACTCAACTGTGGCCAGATCAAGACTGGTCTCCAGCA 1320
 1321 CGTTCCGACCGGTTCGCAAAAGTACAAACAGCTTTCGCGATCGAGAGCTGTTGGCGAC 1380
 1321 CGTTCCGACCGGTTCGCAAAAGTACAAACAGCTTTCGCGATCGAGAGCTGTTGGCGAC 1380
 1381 GCGGCGCTTACGAGGTCGAGCGCATTCACAGCTTTCAGGGCTTAATAAAGCGCTT 1440
 1381 GCGGCGCTTACGAGGTCGAGCGCATTCACAGCTTTCAGGGCTTAATAAAGCGCTT 1440
 1441 TTCCAGCGCCGCTAACCTCAAGGTTTCGCGGCGTTCGTTTCTTACTTCTGCTGT 1500
 1441 TTCCAGCGCCGCTAACCTCAAGGTTTCGCGGCGTTCGTTTCTTACTTCTGCTGT 1500
 1501 GACTATGATCGAGGATTCGCAAGAGAGAGAACTCATAAAGGCTTGTTCCTGCTGT 1560
 1501 GACTATGATCGAGGATTCGCAAGAGAGAGAACTCATAAAGGCTTGTTCCTGCTGT 1560
 1561 CAAGCAGGGAACGTCCTT 1578
 1561 CAAGCAGGGAACGTCCTT 1578

RESULT 3
 US-09-738-626-1
 ; Sequence 1, Application US/09738626
 ; Publication No. US20020197605A1
 ; GENERAL INFORMATION:
 ; APPLICANT: NAKAGAWA, SATOSHI
 ; APPLICANT: MIZOGUCHI, HIROSHI
 ; APPLICANT: ANDO, SEIKO
 ; APPLICANT: HAYASHI, MIKIRO
 ; APPLICANT: OCHIAI, KEIKO
 ; APPLICANT: YOKOI, HARUHIKO
 ; APPLICANT: TATEISHI, NAOKO
 ; APPLICANT: SENOH, AKIHIRO
 ; APPLICANT: IKEDA, MASATO
 ; APPLICANT: OZAKI, AKIO
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
 ; FILE REFERENCE: 249-125
 ; CURRENT APPLICATION NUMBER: US/09/738,626
 ; CURRENT FILING DATE: 2000-12-20
 ; PRIOR APPLICATION NUMBER: JP 99/37484
 ; PRIOR FILING DATE: 1999-12-16
 ; PRIOR APPLICATION NUMBER: JP 00/159162
 ; PRIOR FILING DATE: 2000-04-07
 ; PRIOR APPLICATION NUMBER: JP 00/280988
 ; PRIOR FILING DATE: 2000-08-03
 ; NUMBER OF SEQ ID NOS: 7059
 ; SOFTWARE: Patent in ver. 3.0
 ; SEQ ID NO 1
 ; LENGTH: 3309400
 ; TYPE: DNA
 ; ORGANISM: Corynebacterium glutamicum
 US-09-738-626-1

Query Match 100.0%; Score 1578; DB 9; Length 3309400;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1578; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GCCTGGGATATGGGTAGTGTTCGCACTAATTTCAACTGATGCTCATCGAAACAAGA 60
 DB 1034799 GCCTGGGATATGGGTAGTGTTCGCACTAATTTCAACTGATGCTCATCGAAACAAGA 1034858
 QY 61 TTCTGTCAACAAATTTGGGTAGTGTTCGCACTAATTTCAACTGATGCTCATCGAAACAAGA 120
 DB 1034859 TTCTGTCAACAAATTTGGGTAGTGTTCGCACTAATTTCAACTGATGCTCATCGAAACAAGA 1034918
 QY 121 TTAGTCTCCAAAGTTGGCATAGGAGCCACAGTGGCTGAAATCATGACGATTTCCGCTCGC 180
 DB 1034919 TTAGTCTCCAAAGTTGGCATAGGAGCCACAGTGGCTGAAATCATGACGATTTCCGCTCGC 1034978

No patent data - 1999-12-16
 provided

```
QY 181 GAAATTCGACTCCCGGGTAAACCAACCGTCGAGGAGAGGTTTTCTTGATGACGGT 240
Db 1034979 GAAATTCGACTCCCGGGTAAACCAACCGTCGAGGAGAGGTTTTCTTGATGACGGT 1035038

QY 241 TCCACGGTGTGCGAGGTGTTCCATCCGGCGCATCCACGGGGTCCACGAGGCTCATGAG 300
Db 1035039 TCCACGGTGTGCGAGGTGTTCCATCCGGCGCATCCACGGGGTCCACGAGGCTCATGAG 1035098

QY 301 CTGCGTGACGGTGGCGATCGCTACCTTGGGCAAGGCGGTTTTGAAGGCAAGTTGAAAACGTC 360
Db 1035099 CTGCGTGACGGTGGCGATCGCTACCTTGGGCAAGGCGGTTTTGAAGGCAAGTTGAAAACGTC 1035158

QY 361 AACGAAGAAATCGGCGAAGAGCTCGCTGGCTTAGAGGTGACGATCAACGGCTCATCGAC 420
Db 1035159 AACGAAGAAATCGGCGAAGAGCTCGCTGGCTTAGAGGTGACGATCAACGGCTCATCGAC 1035218

QY 421 GAAGCAATGATCAAGCTTGATGGGCAACCGCCCAACAGTCCCGCTGGGTGCAACCGCAATC 480
Db 1035219 GAAGCAATGATCAAGCTTGATGGGCAACCGCCCAACAGTCCCGCTGGGTGCAACCGCAATC 1035278

QY 481 CTGCGTGTTCATCGCTGTTGCAAAAGCTGCTGCTGATTCGGCAGGCTCCCACTGTTT 540
Db 1035279 CTGCGTGTTCATCGCTGTTGCAAAAGCTGCTGCTGATTCGGCAGGCTCCCACTGTTT 1035338

QY 541 CGCTACATCGGTGGACAAACGACACACGTTTCTTCAGTTCCAAATGATGAACATCATCAAC 600
Db 1035339 CGCTACATCGGTGGACAAACGACACACGTTTCTTCAGTTCCAAATGATGAACATCATCAAC 1035398

QY 601 GGTGGCGCTCAGCTGATCCGGTGTTCAGCTTCAAGGAATTCATGATCGCTCCATCGGT 660
Db 1035399 GGTGGCGCTCAGCTGATCCGGTGTTCAGCTTCAAGGAATTCATGATCGCTCCATCGGT 1035458

QY 661 GCAGAGACTTCTCTGAGGCTCTCCGCAACGGCGGAGGCTTACACGCACTGAAGTCC 720
Db 1035459 GCAGAGACTTCTCTGAGGCTCTCCGCAACGGCGGAGGCTTACACGCACTGAAGTCC 1035518

QY 721 GTCAATCAAGGAAAGGGCTGTGCAACCGGACTTGGCGATGAGGGCGGTTCCGCTCCTTCC 780
Db 1035519 GTCAATCAAGGAAAGGGCTGTGCAACCGGACTTGGCGATGAGGGCGGTTCCGCTCCTTCC 1035578

QY 781 GTGGCTCCACCGGTGAGGCTTGAACCTTATCGTTGAGGCAATCGAAGGCTGGCTTC 840
Db 1035579 GTGGCTCCACCGGTGAGGCTTGAACCTTATCGTTGAGGCAATCGAAGGCTGGCTTC 1035638

QY 841 ACCCGAGCAAGGACATCGCTCTTGCTCTGACAGTTGCTTCTCTGAGTTCTTCAAGGAC 900
Db 1035639 ACCCGAGCAAGGACATCGCTCTTGCTCTGACAGTTGCTTCTCTGAGTTCTTCAAGGAC 1035698

QY 901 GGCACTTACCACTTCGAAGGTGGCGAGCACTCCGAGCTGAGATGGCAAAACGTTTACGCT 960
Db 1035699 GGCACTTACCACTTCGAAGGTGGCGAGCACTCCGAGCTGAGATGGCAAAACGTTTACGCT 1035758

QY 961 GAGCTCGTTGAGCGGTACCCATCGTCTCCATCGAGGACCCACCTGCAGGAAGATGACTGG 1020
Db 1035759 GAGCTCGTTGAGCGGTACCCATCGTCTCCATCGAGGACCCACCTGCAGGAAGATGACTGG 1035818

QY 1021 GAGGTTACACCAACCTCAGCGCAACCATCGGCGACAAGGTTTCAGATCGTTGGCGACGAC 1080
Db 1035819 GAGGTTACACCAACCTCAGCGCAACCATCGGCGACAAGGTTTCAGATCGTTGGCGACGAC 1035878

QY 1081 TTCTTGTGTCACCAACCTTCAGCGGCTTGAAGAGGCGATCGTTAAGAGGCTGCCAATCC 1140
Db 1035879 TTCTTGTGTCACCAACCTTCAGCGGCTTGAAGAGGCGATCGTTAAGAGGCTGCCAATCC 1035938

QY 1141 ATCTCGTTAAGGTGAACCAATCGGTACCTTCAACCGGACCTTCGACGCTGTCGATG 1200
Db 1035939 ATCTCGTTAAGGTGAACCAATCGGTACCTTCAACCGGACCTTCGACGCTGTCGATG 1035998

QY 1201 GCTCACCGCGAGGCTACACCTTCATGATGTCCTCCACCGTTCGGGTGAGACCGAGGACACC 1260
Db 1035999 GCTCACCGCGAGGCTACACCTTCATGATGTCCTCCACCGTTCGGGTGAGACCGAGGACACC 1036058
```

```
QY 1261 ACCATTGCTGACTCGCGAGTTGCATCAACTGTGGCCAGATCAAGACTGTGTGCTCCAGCA 1320
Db 1036059 ACCATTGCTGACTCGCGAGTTGCATCAACTGTGGCCAGATCAAGACTGTGTGCTCCAGCA 1036118

QY 1321 CGTTCCGACCGGTTCGCAAAAGTACACACAGCTTCTCGCATCGAGAGCTGCTTGGCGAC 1380
Db 1036119 CGTTCCGACCGGTTCGCAAAAGTACACACAGCTTCTCGCATCGAGAGCTGCTTGGCGAC 1036178

QY 1381 GCCGGCGTCTACGAGGTCCGAGCGATTCGCCAGCTTTCAGGCTAAAATAAAGCGCTT 1440
Db 1036179 GCCGGCGTCTACGAGGTCCGAGCGATTCGCCAGCTTTCAGGCTAAAATAAAGCGCTT 1036238

QY 1441 TTGCAAGCCCGGTAAACCTCAAGGTTGCGGGCGTGTGCTTACTACTGTACTGTGT 1500
Db 1036239 TTGCAAGCCCGGTAAACCTCAAGGTTGCGGGCGTGTGCTTACTACTGTACTGTGT 1036298

QY 1501 GACTATGATCGAGATTTATGGCAAGGCAAGGCAAGAACTCATAAAGGCTTGTCTCTCT 1560
Db 1036299 GACTATGATCGAGATTTATGGCAAGGCAAGGCAAGAACTCATAAAGGCTTGTCTCTCT 1036358

QY 1561 CAAGCAGGGAACTGCTT 1578
Db 1036359 CAAGCAGGGAACTGCTT 1036376
```

RESULT 4

```
US-09-860-768-3
; Sequence 3, Application US/09860768
; Patent No. US20020082403A1
; GENERAL INFORMATION:
; APPLICANT: Mockel, Bettina
; APPLICANT: Pfeifferle, Walter
; APPLICANT: Hermann, Thomas
; APPLICANT: Pohler, Alfred
; APPLICANT: Kalinowski, Jorn
; APPLICANT: Bathe, Brigitte
; TITLE OF INVENTION: New Nucleotide Sequences that Code for the Eno Gene
; FILE REFERENCE: 21123/278404
; CURRENT APPLICATION NUMBER: US/09/860,768
; CURRENT FILING DATE: 2001-05-21
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patent version 3.0
; SEQ ID NO 3
; LENGTH: 1578
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
US-09-860-768-3
```

```
Query Match 99.9%; Score 1576.4; DB 9; Length 1578;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1577; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 1 GGCTGGGGATATGGGTAGTTTTCGCCACTAAATTTCAACTGATTCCTCATCGAAACAAGA 60
Db 1 GGCTGGGGATATGGGTAGTTTTCGCCACTAAATTTCAACTGATTCCTCATCGAAACAAGA 60

QY 61 TTCTGTGCAACAATTTGGGTGTAGACGTGATTGAAGACATTTTGATCACTGATTAATTTCTAG 120
Db 61 TTCTGTGCAACAATTTGGGTGTAGACGTGATTGAAGACATTTTGATCACTGATTAATTTCTAG 120

QY 121 TTAGCTCCCAAGTTGGCATATAGGAGGCCACAGTGGCTGAAATCATGACGTAATTCGCTCGC 180
Db 121 TTAGCTCCCAAGTTGGCATATAGGAGGCCACAGTGGCTGAAATCATGACGTAATTCGCTCGC 180

QY 181 GAAATTTCTGACTCCCGCGGTAAACCCAAACCGTCGAGGCGAGAGGTTTTCTGGATGACGGT 240
Db 181 GAAATTTCTGACTCCCGCGGTAAACCCAAACCGTCGAGGCGAGAGGTTTTCTGGATGACGGT 240

QY 241 TCCCAAGGCTGTGCGAGGTGTTCCATCCGGCGCATCCACCGGGCGTCCACGAGGCTCATGAG 300
Db 241 TCCCAAGGCTGTGCGAGGTGTTCCATCCGGCGCATCCACCGGGCGTCCACGAGGCTCATGAG 300

QY 301 CTGCGTGACGGTGGCGATCGCTTACCTTCAGTGGCGAAGGCGTTTTGAAGGCAAGTTGAAAACGTC 360
```

Db 301 CTGCGTACGGTGGCGATCGTACCTGGGCAAGGGCGTTTGAAGCGAGTTGAAAACGTC 360
Qy 361 AACGAGAAATCGGCGACGAGCTCGTGGCTAGAGGCTGACGATCAGGCGCTCATCGAC 420
Db 361 AACGAGAAATCGGCGACGAGCTCGTGGCTAGAGGCTGACGATCAGGCGCTCATCGAC 420
Qy 421 GAAGCAATGATCAAGCTTGAATGCGACCGCCCAACAGTCCCGCTCGGGTGCMAACGCAATC 480
Db 421 GAAGCAATGATCAAGCTTGAATGCGACCGCCCAACAGTCCCGCTCGGGTGCMAACGCAATC 480
Qy 481 CTTGGTGTTCATGCTGTGTCAGAGGCTGTGCTGAATTCGCGACGCGCTCCCACTGTTTC 540
Db 481 CTTGGTGTTCATGCTGTGTCAGAGGCTGTGCTGAATTCGCGACGCGCTCCCACTGTTTC 540
Qy 541 CGCTACATCGGTGGGCAACGACACGCTTCTCCAGTTCCTCCAGTTCATGATGATCAATCAAC 600
Db 541 CGCTACATCGGTGGGCAACGACACGCTTCTCCAGTTCCTCCAGTTCATGATGATCAATCAAC 600
Qy 601 GGTGGCGCTCACGCTGACTCCGGTGTGAGCTTTCAGGAATTCATGATCGTCCCAATCGGT 660
Db 601 GGTGGCGCTCACGCTGACTCCGGTGTGAGCTTTCAGGAATTCATGATCGTCCCAATCGGT 660
Qy 661 GCAGAGACTTCTCTGAGGCTCTCCGCAACGGCGGAGGCTTACCAACGCTGAGTCC 720
Db 661 GCAGAGACTTCTCTGAGGCTCTCCGCAACGGCGGAGGCTTACCAACGCTGAGTCC 720
Qy 721 GTCATCAAGAAAGGGCTGTCCACGAGCTTGGCGATGAGGGCGGCTTCGTCCTTCC 780
Db 721 GTCATCAAGAAAGGGCTGTCCACGAGCTTGGCGATGAGGGCGGCTTCGTCCTTCC 780
Qy 781 GTCCGCTCCACCGTGGAGCTCTTACCTTATCGTTGAGGCAATCGAGAAGCTGGCTTC 840
Db 781 GTCCGCTCCACCGTGGAGCTCTTACCTTATCGTTGAGGCAATCGAGAAGCTGGCTTC 840
Qy 841 ACCCGAGGCAAGGACATCGCTCTGCTCTGAGGCTTCTCTCTGAGTTCCTCAAGGAC 900
Db 841 ACCCGAGGCAAGGACATCGCTCTGCTCTGAGGCTTCTCTCTGAGTTCCTCAAGGAC 900
Qy 901 GGCACCTACCTTGAAGGTCGCGACGCTTCCGAGCTGAGATGAGCAACGTTACGCT 960
Db 901 GGCACCTACCTTGAAGGTCGCGACGCTTCCGAGCTGAGATGAGCAACGTTACGCT 960
Qy 961 GAGCTGTGAGCGGTACCAATCGTCTCCATCGAGGACCCACTCGAGCAATGATGACTGG 1020
Db 961 GAGCTGTGAGCGGTACCAATCGTCTCCATCGAGGACCCACTCGAGCAATGATGACTGG 1020
Qy 1021 GAGGGTTACCAACCTCCTCAGCGCAACATCGGCGACAGGTTTCAGATCGTGGCGACGAC 1080
Db 1021 GAGGGTTACCAACCTCCTCAGCGCAACATCGGCGACAGGTTTCAGATCGTGGCGACGAC 1080
Qy 1081 TTCTTCGTCAACACCTGAGCGCTGGAAGGAGGCAATCGTAAAGAGGCTGCCAATCC 1140
Db 1081 TTCTTCGTCAACACCTGAGCGCTGGAAGGAGGCAATCGTAAAGAGGCTGCCAATCC 1140
Qy 1141 ATCTGCTGTTAAGTGAACAGATCGTACCTCAACGAGACCTTCGACGCTGTGACATG 1200
Db 1141 ATCTGCTGTTAAGTGAACAGATCGTACCTCAACGAGACCTTCGACGCTGTGACATG 1200
Qy 1201 GCTCACCGCGAGGCTTACACCTCCATGATGTCCTCCAGCTTCCGCTGAGACCGAGACCC 1260
Db 1201 GCTCACCGCGAGGCTTACACCTCCATGATGTCCTCCAGCTTCCGCTGAGACCGAGACCC 1260
Qy 1261 ACCATTGCTGACCTCGAGTGTGACCTCAACTGTGGCCAGATCAAGACTGGTGTCCAGCA 1320
Db 1261 ACCATTGCTGACCTCGAGTGTGACCTCAACTGTGGCCAGATCAAGACTGGTGTCCAGCA 1320
Qy 1321 CGTTCGACCGGTGCGAAAGTACAAAGCTTCTCCGATCGAGCAGCTGTGTGGCGAC 1380
Db 1321 CGTTCGACCGGTGCGAAAGTACAAAGCTTCTCCGATCGAGCAGCTGTGTGGCGAC 1380
Qy 1381 GCGCGGCTTACGCGAGGTCGAGGCAATTCACAGCTTTCAGGCTTAATTAAGGCTT 1440

Db 1381 GCGCGGCTTACGCGAGGTCGAGGCGCAATTCACAGCTTTCAGGCTTAATTAAGGCTT 1440
Qy 1441 TTCGAGCGCGGTACCTCAAGGTTGCGGGCGTTCGCTTACTCTGTTACTGGTCT 1500
Db 1441 TTCGAGCGCGGTACCTCAAGGTTGCGGGCGTTCGCTTACTCTGTTACTGGTCT 1500
Qy 1501 GACTATGATCGAGGATTTATGCAAGCAGAGAAACTCATAAAGCGCTTGTTCCTGTCT 1560
Db 1501 GACTATGATCGAGGATTTATGCAAGCAGAGAAACTCATAAAGCGCTTGTTCCTGTCT 1560
Qy 1561 CAAGCAGGGAACGTGCTT 1578
Db 1561 CAAGCAGGGAACGTGCTT 1578

RESULT 5
US-10-728-947-3
; Sequence 3, Application US/10728947
; Publication No. US20040220394A1
; GENERAL INFORMATION:
; APPLICANT: Mockel, Bettina
; APPLICANT: Pfeifferle, Walter
; APPLICANT: Hermann, Thomas
; APPLICANT: Pohler, Alfred
; APPLICANT: Kalinowski, Jorn
; APPLICANT: Bathe, Brigitte
; TITLE OF INVENTION: New Nucleotide Sequences that Code for the Eno Gene
; FILE REFERENCE: 21123/278404
; CURRENT APPLICATION NUMBER: US/10/728,947
; CURRENT FILING DATE: 2003-12-08
; PRIOR APPLICATION NUMBER: US/09/860,768
; PRIOR FILING DATE: 2001-05-21
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 3
; LENGTH: 1578
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
US-10-728-947-3

Query Match 99.9%; Score 1576.4; DB 20; Length 1578;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1577; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCGTGGGATATGGTAGTATTTTCGCCACTAATTTCAACTGATTCCTCATCGAAACAAGA 60
Db 1 GCGTGGGATATGGTAGTATTTTCGCCACTAATTTTCGCCACTAATTTCAACTGATTCCTCATCGAAACAAGA 60
Qy 61 TTCGTGCAACAATTTGGGTGTAGACGTGATTGGAAGACATTTTGATCACGTGAATAATTCCTAG 120
Db 61 TTCGTGCAACAATTTGGGTGTAGACGTGATTGGAAGACATTTTGATCACGTGAATAATTCCTAG 120
Qy 121 TTAGCTCCCAAGTTTGGCATAGGAGGCCACAGTGGCTGAAATCATGACGTATTCCTCGC 180
Db 121 TTAGCTCCCAAGTTTGGCATAGGAGGCCACAGTGGCTGAAATCATGACGTATTCCTCGC 180
Qy 181 GAAATTCCTGACTCCCGCGGTAAACCAACCGTTCGAGGAGGAGGTTTTCCTGATGACCGT 240
Db 181 GAAATTCCTGACTCCCGCGGTAAACCAACCGTTCGAGGAGGAGGTTTTCCTGATGACCGT 240
Qy 241 TCCCAAGGTTGCGAGGTTTCATTCGCGCGCATCCACCGGCTCCAGAGGCTCATGAG 300
Db 241 TCCCAAGGTTGCGAGGTTTCATTCGCGCGCATCCACCGGCTCCAGAGGCTCATGAG 300
Qy 301 CTGCGTGAACGTTGGCGATCGCTACTCGGGCAAGGCGGTTTGAAGGAGTTGAAAACGTC 360
Db 301 CTGCGTGAACGTTGGCGATCGCTACTCGGGCAAGGCGGTTTGAAGGAGTTGAAAACGTC 360
Qy 361 AACGAGAAATTCGCGACGAGCTCGTGGCTTAGAGGCTGACGATCAGCGCTCATCGAC 420
Db 361 AACGAGAAATTCGCGACGAGCTCGTGGCTTAGAGGCTGACGATCAGCGCTCATCGAC 420
Qy 421 GAAGCAATGATCAAGCTTGAATGCGACCGCCCAACAGTCCCGCTCGGGTGCMAACGCAATC 480

|||||
421 GAAGCAATGATCAAGCTTATGCGCAACCGCAACAAAGTCCCGCTGGGTGCAACGCAATC 480
QY
481 CTTGGTGTTCATGGCTGTTGCAAAAGCTGCTGCTGATTCGGCAGGCTCCCACTGTTT 540
Db
481 CTTGGTGTTCATGGCTGTTGCAAAAGCTGCTGCTGATTCGGCAGGCTCCCACTGTTT 540
QY
541 CGCTACATCGGTGGACCAAAAGCACAAGTCTTCTTCAGTTCCMAATGATGAATCATCAAC 600
Db
541 CGCTACATCGGTGGACCAAAAGCACAAGTCTTCTTCAGTTCCMAATGATGAATCATCAAC 600
QY
601 GGTGGCGCTCAGCGTACTCCGGTGTTCAGGTTTCAAGGAATTCATGATCGTCCCAATCGGT 660
Db
601 GGTGGCGCTCAGCGTACTCCGGTGTTCAGGTTTCAAGGAATTCATGATCGTCCCAATCGGT 660
QY
661 GCAGAGACCTTCTCTGAGGCTCTCCGCAACGGCGCGAGGTCACACGCACTGAAGTCC 720
Db
661 GCAGAGACCTTCTCTGAGGCTCTCCGCAACGGCGCGAGGTCACACGCACTGAAGTCC 720
QY
721 GTCAATCAAGAAAGGGCTGTCCACCGGACTTTGGCGATGAGGGCGGCTTCGCTCCTTCC 780
Db
721 GTCAATCAAGAAAGGGCTGTCCACCGGACTTTGGCGATGAGGGCGGCTTCGCTCCTTCC 780
QY
781 GTCGGCTCCACCGGTGAGGCTCTTGACCTTATCGTTGAGGCAATCGAAGGCTGGCTTC 840
Db
781 GTCGGCTCCACCGGTGAGGCTCTTGACCTTATCGTTGAGGCAATCGAAGGCTGGCTTC 840
QY
841 ACCCAGGCAAGGACATCGCTCTTGCTCTGGACGTTGCTCTCTGAGTTCTTCAAGGAC 900
Db
841 ACCCAGGCAAGGACATCGCTCTTGCTCTGGACGTTGCTCTCTGAGTTCTTCAAGGAC 900
QY
901 GGCACCTACCACTTCGAAGGTGGCGAGCACTCCGGCAGCTGAGATGGCAAAAGTTACGCT 960
Db
901 GGCACCTACCACTTCGAAGGTGGCGAGCACTCCGGCAGCTGAGATGGCAAAAGTTACGCT 960
QY
961 GAGCTCGTTGACGGGTACCCAACTCGTCTCCATCGAGGACCCACTGCGAGAGATGACTGG 1020
Db
961 GAGCTCGTTGACGGGTACCCAACTCGTCTCCATCGAGGACCCACTGCGAGAGATGACTGG 1020
QY
1021 GAGGTTTACACCAACCTCACCGCAACCATCGCGCAAGGTTTCAGATCGTTGGCGAGAC 1080
Db
1021 GAGGTTTACACCAACCTCACCGCAACCATCGCGCAAGGTTTCAGATCGTTGGCGAGAC 1080
QY
1081 TTCTTGTTCACCAACCTTCAGAGGCTGTAAGAGGGCATCGCTAAGAGGCTGCGCAACTCC 1140
Db
1081 TTCTTGTTCACCAACCTTCAGAGGCTGTAAGAGGGCATCGCTAAGAGGCTGCGCAACTCC 1140
QY
1141 ATCTGTTAAGGTGAACAGATCGGTACCTCTCACCGAGACCTTCGACGCTGTCGACATG 1200
Db
1141 ATCTGTTAAGGTGAACAGATCGGTACCTCTCACCGAGACCTTCGACGCTGTCGACATG 1200
QY
1201 GCTCACCGCGCAGGCTACACCTCCATGATGCCCAACCGTTCCGGTGGAGACCGAGGACAC 1260
Db
1201 GCTCACCGCGCAGGCTACACCTCCATGATGCCCAACCGTTCCGGTGGAGACCGAGGACAC 1260
QY
1261 ACCATTGCTGACCTCGCAGTTGGCATCTAACTGTGGCCAGATCAAGACTGGTGTCCAGCA 1320
Db
1261 ACCATTGCTGACCTCGCAGTTGGCATCTAACTGTGGCCAGATCAAGACTGGTGTCCAGCA 1320
QY
1321 GGTTCGACCGGTGCGCAAGTACAACAGGCTTCTCCGCACTCGAGGAGCTGCTTGGCGAC 1380
Db
1321 GGTTCGACCGGTGCGCAAGTACAACAGGCTTCTCCGCACTCGAGGAGCTGCTTGGCGAC 1380
QY
1381 GCCGGGCTCTACGAGGTGCGAGGCAATCCACGCTTTTCAGGGCTAAATAAAGCGCTT 1440
Db
1381 GCCGGGCTCTACGAGGTGCGAGGCAATCCACGCTTTTCAGGGCTAAATAAAGCGCTT 1440
QY
1441 TTGACCGCGGTAACTCAAGTTTGGCGGCTGCTTGCCTTACTACTGTTACTGTTGT 1500
Db
1441 TTGACCGCGGTAACTCAAGTTTGGCGGCTGCTTGCCTTACTACTGTTACTGTTGT 1500
QY
1501 GACTATGATCGAGGATTAAGCAAGCAGAGAAAGCTCATAAAGGCTGCTGCTGCTCT 1560
|||||

Db 1501 GACTATGATCGAGGATTAAGCAAGCAGAGAAAGCTCATAAAGGCTGCTGCTGCTCT 1560
QY 1561 CAAGCAGGAAACGTGCTT 1578
Db 1561 CAAGCAGGAAACGTGCTT 1578
RESULT 6
US-10-494-836-7
; Sequence 7, Application US/10494836
; Publication No. US20050014233A1
; GENERAL INFORMATION:
; APPLICANT: Zeider, Oskar
; APPLICANT: Pompejus, Markus
; APPLICANT: Schroder, Burkhard
; APPLICANT: Kroger, Burkhard
; APPLICANT: Klopprogge, Corinna
; APPLICANT: Haberhauser, Gregor
; TITLE OF INVENTION: Genes coding for proteins of carbon metabolism and energy produc
; FILE REFERENCE: BGI-167US
; CURRENT APPLICATION NUMBER: US/10/494,836
; PRIOR FILING DATE: 2004-05-05
; PRIOR FILING DATE: 2002-10-31
; PRIOR FILING DATE: 2002-10-31
; PRIOR FILING DATE: 2002-10-31
; PRIOR FILING DATE: 2001-11-05
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 7
; LENGTH: 1405
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (101)..(1375)
; OTHER INFORMATION: RXA00235
US-10-494-836-7

Query Match 88.9%; Score 1403.4; DB 21; Length 1405;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1404; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 51 CGAAACAAGATTCTGCGCAACAAATGGGTGTAGACGTGATGAAGACATTTGATCACGTGA 110
Db 1 CGAAACAAGATTCTGCGCAACAAATGGGTGTAGACGTGATGAAGACATTTGATCACGTGA 60
QY 111 ATAAATCTAGTTAGTCTCCCAAGTTGCTAGAGGCGCACAGTGGCTGAAATCATGACGT 170
Db 61 ATAAATCTAGTTAGTCTCCCAAGTTGCTAGAGGCGCACAGTGGCTGAAATCATGACGT 120
QY 171 ATTCGCTCGCAAAATTTCTGACTCCCGCGGTAAACCAACCGTCGAGGAGAGGTTTTCTT 230
Db 121 ATTCGCTCGCAAAATTTCTGACTCCCGCGGTAAACCAACCGTCGAGGAGAGGTTTTCTT 180
QY 231 GGATGACGTTTCCCAAGGTTGCGAGGTTTCCATCGCGGCGATCCACCGGCGTCCACGA 290
Db 181 GGATGACGTTTCCCAAGGTTGCGAGGTTTCCATCGCGGCGATCCACCGGCGTCCACGA 240
QY 291 GGCTCATGAGCTCGGTGACGCGTGGCGATCGTACCTGGGCAAGGGCGTTTTGAAGGCGGT 350
Db 241 GGCTCATGAGCTCGGTGACGCGTGGCGATCGTACCTGGGCAAGGGCGTTTTGAAGGCGGT 300
QY 351 TGAACAAGTCAACGAAGAAATCGGCGACGAGCTCGCTGGGCTTAGAGGCTGACGATCAGCG 410
Db 301 TGAACAAGTCAACGAAGAAATCGGCGACGAGCTCGCTGGGCTTAGAGGCTGACGATCAGCG 360
QY 411 CCTCATCGAGGAGCAATGATCAAGTTTATGCGACCGCGCAACAAAGTCCCGGCTGGGTGC 470
Db 361 CCTCATCGAGGAGCAATGATCAAGTTTATGCGACCGCGCAACAAAGTCCCGGCTGGGTGC 420
QY 471 AAACGCAATCTTGGTGTTCATGAGTGTTCGAAAGGCTGCTGCTGATTTCCGACGCGCT 530
Db 421 AAACGCAATCTTGGTGTTCATGAGTGTTCGAAAGGCTGCTGCTGATTTCCGACGCGCT 480

RESULT 7
US-10-781-014-71
; Sequence 71: Application US/10781014
; Publication No. US20040180408A1
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schroder, Hartwig

Query Match	88.5%;	Score 1396.4;	DB 19;	Length 1398;
Best Local Similarity	99.9%;	Pred. No. 0;		
Matches 1397;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;
51	CGAAACAAGATTCTGTCGCAACAATTTGGGTGTAGACGCTGATTGAAGACATTTTGATCACTGA	110		
1	CGAAACAAGATTCTGTCGCAACAATTTGGGTGTAGACGCTGATTGAAGACATTTTGATCACTGA	60		
111	ATAATTCCTAGTTAGTCTCCCAAGTTGGCATPAGGAGGCCACAGTGGCTGGAATCATGCACGT	170		
61	ATAATTCCTAGTTAGTCTCCCAAGTTGGCATPAGGAGGCCACAGTGGCTGGAATCATGCACGT	120		
171	ATTTCGCTCGCGAAATTTCTCGACTCCCGCGGTAAACCAACCGTCGAGGCGAGAGGTTTTCCT	230		
121	ATTTCGCTCGCGAAATTTCTCGACTCCCGCGGTAAACCAACCGTCGAGGCGAGAGGTTTTCCT	180		
231	GGATGACGGTTCCCAACGGTGTCCGAGGTGTTCCATCCGGCGCATCCACGGGCTCCACGA	290		
181	GGATGACGGTTCCCAACGGTGTCCGAGGTGTTCCATCCGGCGCATCCACGGGCTCCACGA	240		
291	GGCTCATAGAGCTGCGTGCAGGTGGCGCATCGCTACCTGGGCAAGGGGCTTTTGAAGCAGT	350		
241	GGCTCATAGAGCTGCGTGCAGGTGGCGCATCGCTACCTGGGCAAGGGGCTTTTGAAGCAGT	300		
351	TGAAAACGCTCAACGAAGAAATCGGCGACGAGCTCGCTGSCCTAGAGGCTGACGATCAGCG	410		
301	TGAAAACGCTCAACGAAGAAATCGGCGACGAGCTCGCTGSCCTAGAGGCTGACGATCAGCG	360		
411	CCTCATCGACGAGCAATGATCAAGCTTGATGGCACCGCCACAAGTCCCGCTGGGTGC	470		
361	CCTCATCGACGAGCAATGATCAAGCTTGATGGCACCGCCACAAGTCCCGCTGGGTGC	420		
471	AAACGCAATTCCTTGGTGTTCCTATCGGCTGTTGCCAAGGCTGCTGCTGATTCCGCAAGGCT	530		


```

421 AAACCAATCCCTGGTGTTCATGGCTGTGCAAAAGGCTGCTGCTGATCCGAGGCT 480
531 CCCACTGTTCCGCTACATCGGTGGACCAAAAGGCAACAGCTTCTCCAGTTCCTCAATGATGAA 590
481 CCCACTGTTCCGCTACATCGGTGGACCAAAAGGCAACAGCTTCTCCAGTTCCTCAATGATGAA 540
591 CATCATCAACCGGTGGCTGCTACGCTGATCTCCGCTGTTGACGTTCCAGGAATTCATGATCGC 650
541 CATCATCAACCGGTGGCTGCTACGCTGATCTCCGCTGTTGACGTTCCAGGAATTCATGATCGC 600
651 TCCCAATCGGTGAGAGAGCTTCTGAGGCTCTCCGCAACCGCGGAGGCTTACCAAGC 710
601 TCCCAATCGGTGAGAGAGCTTCTGAGGCTCTCCGCAACCGCGGAGGCTTACCAAGC 660
711 ACTGAAGTCCGCTCATCAAGGAAGAGGCTCTGTCACCGGACTTGGCGATGAGGCGGCTT 770
661 ACTGAAGTCCGCTCATCAAGGAAGAGGCTCTGTCACCGGACTTGGCGATGAGGCGGCTT 720
771 CGCTCTTCCGCTGAGCTTCCACCGGTGAGGCTTCTGACCTTATCGTTGAGGCAATTCGAGAA 830
721 CGCTCTTCCGCTGAGCTTCCACCGGTGAGGCTTCTGACCTTATCGTTGAGGCAATTCGAGAA 780
831 GGCTGGCTTACCCGAGGCAAGGACATCGCTCTTGTCTGACGCTTCTTCTGAGTT 890
781 GGCTGGCTTACCCGAGGCAAGGACATCGCTCTTGTCTGACGCTTCTTCTGAGTT 840
891 CTTCAAGAGCGGCACTTACCACTTCAAGAGTGGCCAGCACTCCGAGCTGAGATGGCAAA 950
841 CTTCAAGAGCGGCACTTACCACTTCAAGAGTGGCCAGCACTCCGAGCTGAGATGGCAAA 900
951 CGTTTACGCTGAGCTGCTGAGGCTTCCACCGGTGAGGCTTCTGACCTTATCGTTGAGGCAATTCGAGAA 1010
901 CGTTTACGCTGAGCTGCTGAGGCTTCCACCGGTGAGGCTTCTGACCTTATCGTTGAGGCAATTCGAGAA 960
1011 AGATGATCGGAGGAGGTATACCAACCTTCAAGAGTGGCCAGCACTCCGAGCTGAGATGGCAAA 1070
961 AGATGATCGGAGGAGGTATACCAACCTTCAAGAGTGGCCAGCACTCCGAGCTGAGATGGCAAA 1020
1071 TGGCGAGCACTTCTTGTGACCAACCTTCAAGAGTGGCCAGCACTCCGAGCTGAGATGGCAAA 1130
1021 TGGCGAGCACTTCTTGTGACCAACCTTCAAGAGTGGCCAGCACTCCGAGCTGAGATGGCAAA 1080
1131 TGGCAACTCCATCTGTTTAAAGGTGAACAGATCGGTACCTTCAACCGAGACCTTTCGAGCG 1190
1081 TGGCAACTCCATCTGTTTAAAGGTGAACAGATCGGTACCTTCAACCGAGACCTTTCGAGCG 1140
1191 TGTGCAATGGCTCACCGCGAGGCTACACCTCCATGATGTCCACCGTTCGCGTGAGAC 1250
1141 TGTGCAATGGCTCACCGCGAGGCTACACCTCCATGATGTCCACCGTTCGCGTGAGAC 1200
1251 CGAGGACACCACTTGTGACCTCGAGTGTGCACTCAACTGTGGCGAGATCAAGACTGG 1310
1201 CGAGGACACCACTTGTGACCTCGAGTGTGCACTCAACTGTGGCGAGATCAAGACTGG 1260
1311 TGTCTCGAGCAGTTCGACCGGTGTCGCAAGATCAACAGCTTCTCCGATTCGAGCAGCT 1370
1261 TGTCTCGAGCAGTTCGACCGGTGTCGCAAGATCAACAGCTTCTCCGATTCGAGCAGCT 1320
1371 GCTTGGCGAGCGCGGCTTACGAGGTTCGAGCGGATTCCTCAGCTTTCAGGGCTAAAT 1430
1321 GCTTGGCGAGCGCGGCTTACGAGGTTCGAGCGGATTCCTCAGCTTTCAGGGCTAAAT 1380
1431 AAAAGCGCTTTCGAGCG 1448
1381 AAAAGCGCTTTCGAGCG 1398

```

RESULT 8
 US-09-738-626-1085
 ; Sequence 1085, Application US/09738626
 ; Publication No. US20020197605A1
 ; GENERAL INFORMATION:

```

; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SBIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09-738-626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 1085
; LENGTH: 1275
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; US-09-738-626-1085

```

```

Query Match      80.8%; Score 1275; DB 9; Length 1275;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1275; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 151 GTGGCTGAAATCATGACGATTTGCTCGCGAAATTCGACTCCCGCGGTAAACCAACC 210
DB 1 GTGGCTGAAATCATGACGATTTGCTCGCGAAATTCGACTCCCGCGGTAAACCAACC 60

QY 211 GTCGAGCGAGAGGTTTCTCGATGACGTTTCCGAGTTCGCGAGTTCGATCCCGC 270
DB 61 GTCGAGCGAGAGGTTTCTCGATGACGTTTCCGAGTTCGCGAGTTCGATCCCGC 120

QY 271 GCATCCACCGCGCTCCACGAGCTCATGAGTTCGCTGAGCGGTGCGATCGCTACCTGGC 330
DB 121 GCATCCACCGCGCTCCACGAGCTCATGAGTTCGCTGAGCGGTGCGATCGCTACCTGGC 180

QY 331 AAGGGCGTTTGAAGCGAGTTGAAACGTCACAGAGAAATCGGCGACGAGCTCGCTGCG 390
DB 181 AAGGGCGTTTGAAGCGAGTTGAAACGTCACAGAGAAATCGGCGACGAGCTCGCTGCG 240

QY 391 CTAGAGGCTGACGATCAGCGCTCATCGAGAGCAATGATCAAGTTGATGGCACCGCC 450
DB 241 CTAGAGGCTGACGATCAGCGCTCATCGAGAGCAATGATCAAGTTGATGGCACCGCC 300

QY 451 AACAAAGTCCCGCTGGGTGCAACCAATCTTGTGTTTCCATGGCTGTTGCAAGGCT 510
DB 301 AACAAAGTCCCGCTGGGTGCAACCAATCTTGTGTTTCCATGGCTGTTGCAAGGCT 360

QY 511 GCTGCTGATTCGCGAGGCTCCCACTGTTTCCGCTACATCGGTGAGCAACCAACGACAGCT 570
DB 361 GCTGCTGATTCGCGAGGCTCCCACTGTTTCCGCTACATCGGTGAGCAACCAACGACAGCT 420

QY 571 CTTCCAGTTCCCAATGATGATCAACCGGTGCGCTCAGCTGATCCCGTGTGTGAC 630
DB 421 CTTCCAGTTCCCAATGATGATCAACCGGTGCGCTCAGCTGATCCCGTGTGTGAC 480

QY 631 GTTCAGGAATTCATGATCGCTCAATCGGTGAGAGACCTTCTCTGAGGCTCTCCGCAAC 690
DB 481 GTTCAGGAATTCATGATCGCTCAATCGGTGAGAGACCTTCTCTGAGGCTCTCCGCAAC 540

QY 691 GCGCGGAGGCTTCAACGAGCTGAGTTCGCTCATCAAGAGAAAGGGCTGTGCCACCGGA 750
DB 541 GCGCGGAGGCTTCAACGAGCTGAGTTCGCTCATCAAGAGAAAGGGCTGTGCCACCGGA 600

QY 751 CTTGGCGATGAGGGGCGGCTTCCGCTCCTCCGCTCGGCTCCACCGGTGAGGCTCTTGACCTT 810

```



```
Db 601 CTTGGCGATGAGGCGCTTCGCTCTCTCCGTCGGCTCAACCGCTGAGGCTCTTGACCTT 660
Qy 811 ATCGTTGAGCAATCGAGAAGGCTGCTTCAACCCAGGCAAGGACATCGCTCTTGCTCTG 870
Db 661 ATCGTTGAGCAATCGAGAAGGCTGCTTCAACCCAGGCAAGGACATCGCTCTTGCTCTG 720
Qy 871 GACGTTGCTTCTCTGAGTTCTTCAAGGACGCGACCTTACACCTTCAAGGTGCGCAGCAC 930
Db 721 GACGTTGCTTCTCTGAGTTCTTCAAGGACGCGACCTTACACCTTCAAGGTGCGCAGCAC 780
Qy 931 TCGCAGCTGAGATGCAACAGCTTTTACGCTGAGCTGTTGACCGTACCCTATCGTCTCC 990
Db 781 TCGCAGCTGAGATGCAACAGCTTTTACGCTGAGCTGTTGACCGTACCCTATCGTCTCC 840
Qy 991 ATCGAGGACCCACTCGCAGGAAGATGACTGGAGGGTTACACCAACTCACCGCAACCATC 1050
Db 841 ATCGAGGACCCACTCGCAGGAGATGACTGGAGGGTTACACCAACTCACCGCAACCATC 900
Qy 1051 GCGCAAGAGTTTCAGATCGTTGGCGACGACTTCTTCTGTCACCAACCTTCAAGGTGCGCAG 1110
Db 901 GCGCAAGAGTTTCAGATCGTTGGCGACGACTTCTTCTGTCACCAACCTTCAAGGTGCGCAG 960
Qy 1111 GAGGGCATCGTAAGAGGCTGCAACTCTGCTTAAAGTGAACAGATCGGTACC 1170
Db 961 GAGGGCATCGTAAGAGGCTGCAACTCTGCTTAAAGTGAACAGATCGGTACC 1020
Qy 1171 CTCACCGACCTTCGACGCTGTCGACATGGCTCACCGCGCAGGCTACACCTTCCATGATG 1230
Db 1021 CTCACCGACCTTCGACGCTGTCGACATGGCTCACCGCGCAGGCTACACCTTCCATGATG 1080
Qy 1231 TCCACCGCTTCGGGTGAGACCGAGGACACCACTTGTGCTGACCTCGCAGTTGCACTCAAC 1290
Db 1081 TCCACCGCTTCGGGTGAGACCGAGGACACCACTTGTGCTGACCTCGCAGTTGCACTCAAC 1140
Qy 1291 TGTGGCAGATCAAGCTGCTGTCAGACGCTTCCGACCGGTGTCGCAAGGTACAAACGAG 1350
Db 1141 TGTGGCAGATCAAGCTGCTGTCAGACGCTTCCGACCGGTGTCGCAAGGTACAAACGAG 1200
Qy 1351 CTTCTCCGATCGACGAGCTGCTTGGCGACGCGGCTTACGCGAGGTCGACGCGCATTC 1410
Db 1201 CTTCTCCGATCGACGAGCTGCTTGGCGACGCGGCTTACGCGAGGTCGACGCGCATTC 1260
Qy 1411 CCACGCTTTCAGGCG 1425
Db 1261 CCACGCTTTCAGGCG 1275
```

RESULT 9

```
US-10-728-122A-17724
; Sequence 17724, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangueu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Cart, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
```

```
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 17724
; LENGTH: 1275
; TYPE: DNA
; ORGANISM: Corynebacterium diptheriae
US-10-282-122A-17724
```

Query Match

Best Local Similarity 57.6%; Score 908.6; DB 17; Length 1275;

Matches 1046; Conservative 0; Mismatches 229; Indels 0; Gaps 0;

```
Qy 151 GTGCTGAAATCATGCACGATTCGCTCGCGAAATTCGACTCCCGCGTAACCAACC 210
Db 1 GTGCTGACATTATGCAGTATTGCTGCTGTAATCTTGGACTCCCGCGTAACCAACC 60
Qy 211 GTCGAGGACAGAGTTTCTCGATGACGCTTCCACGCTGCGAGGTTCATCCGCG 270
Db 61 GTTGAAGCCGAAGTTTCTTGGATGACGATCCACGCGTTGCGGCGTTCCTTCCGCT 120
Qy 271 GCATCCACCGCGCTCCAGGCTCATGAGCTGCTGAGCGGTGCGAGTACGCTACCTGGCG 330
Db 121 GCTCCACCGGTGTTTACGAGGCTTCCAGGCTTCCGAGCTTGGAGCGGTACCTCGCG 180
Qy 331 AAGGCGGTTTTGAAGGCGATTTGAAACGTCACGAGAAATTCGCGACGAGCTCGCTGCG 390
Db 181 AAGGCGGTTCTCAATGCAAGTTTAAACGTAACGAGGAAATCGCTGACGCAATCGTGT 240
Qy 391 CTAGAGGCTGAGATCAGCGCTCATCGACGAGCAATGATCAAGCTTGTATGCGACCGCC 450
Db 241 GCAGAAGCCGACGATCAGCGCTTATCGATCAGGCGATGATTGCTCTCGACGCACTGAG 300
Qy 451 AACAGTCCCGCTGGGTGCAACGCAATCCTTGGTGTTCATGCTGCTTTCGCAAGGCT 510
Db 301 AACAGTCTCGTCTCGGCGCAACGCTATCTTGGGTGTATCCATCGCGGTAGTAAAGCT 360
Qy 511 GCTGCTGATTCGCGAGGCTCCCACTGTTCGCTTACATCGGTGACCAAGGCGCACACGCT 570
Db 361 GCGCAGAGTCTGCTGGCTTACCTTTGTACGCTTACATCGGCGGCTTAAAGCTTACGCT 420
Qy 571 CTTCCAGTTTCAATGATGAACATCATCAACGCTGGCGCTACGCTGATTCGCGGTGTTGAC 630
Db 421 CTTCCAGTTTCAATGATGAACATTTGTAACGCTGGCGCACACGCTGACTCCGCGGTTGAT 480
Qy 631 GTTCAGGAATTCATGATCGCTTCCATCGGTGACAGACCTTCTCTGAGGCTCTCGGCAAC 690
Db 481 GTTCAGGAGTTTCAATGATGCTTCTATCGGTGCGAGTCTTCTCTGAGGCTCTCGCATG 540
Qy 691 GCGCGGAGGCTTACACGCACTCAAGTTCGCTATCAAGGAAAGGCGCTCTGCAACGGA 750
Db 541 GGTGCAAGGTTTACCACTCTTGAAGTTCGCTGATTAAGTCAAGGAACTTTTCCACCGCG 600
Qy 751 CTTGGCGATGAGGCGGCTTTCGCTTCTTCCGTGCGCTCCACCGGTGAGGCTTTCGACCT 810
Db 601 CTCGCGCAGGAAGTGTTTCGCACTTCTGTTGAGTCCACCAAGGCGAGCTCTCGACCTC 660
```

811 ATCGTTGAGGCAATCGAAGGCTGGCTTACCCCGAGGAGGACATCGCTCTTGGCTG 870
 661 ATCGTTGAGGCAATCGAAGGCTGGCTTACCCCGAGGAGGACATCGCTCTTGGCTG 720
 871 GACGTTGCTCTCTGAGTCTTCAAGACGCGACCTACCACTTCGAAGGTGGCCAGCAC 930
 721 GACGTTGCTCTCTGAGTCTTCAAGAGTGGCAAGTACCACTTCGAGGGGGGCGAGC 780
 931 TCCGAGCTGAGATGGCAACGTTTACGCTGAGCTCGTTGACGCGTACCCCAATCGTCTCC 990
 781 ACCGCTGAGGAGATGGCAAGTCTACGAGCAGCTCATCGCTGAGTACCCCAATGTTTCC 840
 991 ATCGAGGACCACTGACGAGAGATGACTGGGAGGTTACCAACCTTACCGCAACCAATC 1050
 841 ATCGAGGACCACTGACGAGAGATGACTGGGAGGTTACCAACCTTACCGCAACCAATC 900
 1051 GCGCAAGGTTTCAGATCGTTGGGACGACTTCTTCGTCACCAACCTTGGCGCTGAAG 1110
 901 GGTGACCAAGTTTCAGATCGTGGGACGACTTCTTCGTCACCAACCTTGGCGCTGAAG 960
 1111 GAGGCACTGCTAAGAGGCTGCCAATCTTCCTGTTAAGGTGAACAGATCGGTACC 1170
 961 GAAGGCACTGAGAGAGGCTGCCAATCTTCCTGTTAAGGTGAACAGATCGGTACC 1020
 1171 CTCACCGAGACTTCGACGCTGTGACATGGCTCACCGCGCAGGCTACACCTCCATGATG 1230
 1021 CTGACCGAGACTTCGACGCTGTGATCTCGCACACCGCGCTACCGCACCATGATG 1080
 1231 TCCACCGCTTCGGTGACAGCGAGACACCACTTCGCTGACCACTTCGCTGCAAGTTCGACTCAAC 1290
 1081 TCCACCGCTTCGGTGACAGCGAGACACCACTTCGCTGACCACTTCGCTGCAAGTTCGACTCAAC 1140
 1291 TGTGCGCAGATCAAGACTGTGCTCCAGCAGCTTCGAGCCGCTGTCGCAAGTACCAACAG 1350
 1141 TGTGCGCAGATCAAGACTGTGCTCCAGCAGCTTCGAGCCGCTGTCGCAAGTACCAACAG 1200
 1351 CTTCTCCGATCGAGCAGCTGTTGGCGACGCGCGCGCTCTACGAGGTTCGAGCGCATTC 1410
 1201 CTTCTCCGATCGAGCAGCTGTTGACGATGACGAGTCTACGAGCGCTGTCGAGCGCTTCAGCATTT 1260
 1411 CCAGCTTTCAGGGC 1425
 1261 CCAGCTTTCAGGGC 1275

RESULT 10

US-10-156-761-1/c
 ; Sequence 1, Application US/10156761
 ; Publication No. US20030119018A1
 ; GENERAL INFORMATION:
 ; APPLICANT: OMURA, SATOSHI
 ; APPLICANT: IKEDA, HARUO
 ; APPLICANT: ISHIKAWA, JUN
 ; APPLICANT: HORIKAWA, HIROSHI
 ; APPLICANT: SHIBA, TADAYOSHI
 ; APPLICANT: SAKAKI, YOSHIYUKI
 ; APPLICANT: HATTORI, MASAHIRA
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
 ; FILE REFERENCE: 249-262
 ; CURRENT APPLICATION NUMBER: US/10/156,761
 ; CURRENT FILING DATE: 2002-05-29
 ; PRIOR APPLICATION NUMBER: JP 2001-204089
 ; PRIOR FILING DATE: 2001-05-30
 ; PRIOR APPLICATION NUMBER: JP 2001-272697
 ; PRIOR FILING DATE: 2001-08-02
 ; NUMBER OF SEQ ID NOS: 15109
 ; SEQ ID NO 1
 ; LENGTH: 9025608
 ; TYPE: DNA
 ; ORGANISM: Streptomyces avermitilis
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (4187715)

OTHER INFORMATION: a, t, c, g, other or unknown
 US-10-156-761-1
 Query Match 44.3%; Score 698.4; DB 15; Length 9025608;
 Best Local Similarity 72.0%; Pred. No. 1.1e-202;
 Matches 925; Conservative 0; Mismatches 356; Indels 3; Gaps 1;
 QY 140 AGGAGGCCACAGTGGCTGAAATCATGACATATTCGCTCGGAAATTTCTGACTCCGCG 199
 DB 4373549 AGGAGATGCTCGTCCGCTCCATCGACGTCGTCGAGCCGGAAATCTCGACTCCGAG 4373490
 QY 200 GTAACCCAAACCGTCGAGGACAGGTTTCTCGATGACGCTTCCACGCTGTCGACGCTG 259
 DB 4373489 GCAACCCCAACCGTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTC 4373430
 QY 260 TTCCATCCGCGCATCCACCGCGCTCCACGAGGCTCATGAGCTCGCTGAGCGTGGTGGTGG 316
 DB 4373429 TCCGCTCCGCGCTCCACCGGTCCTTCGAGGCCATCGAGCTCCGCGCAGCGTGACCCCA 4373370
 QY 317 ATCGCTACCTGGGCAAGGCGTTTGAAGCAGTTGAAACGTTGAAACGTTGAAACGTTGAAACGTT 376
 DB 4373369 ACCGTTACCAAGGCAAGGCTTCGAGAAAGCGCTTCCTCGCGCTCATCGACGATCGGCC 4373310
 QY 377 ACAGCTCGCTGGCTAGAGCTGACGATCAGCGCTCATCGAGAGCAATGATCAAGC 436
 DB 4373309 CGAGCTCGCTGGCTAGAGCTGACGATCAGCGCTCATCGAGAGCAATGATCAAGC 4373250
 QY 437 TTGATGACACCGCAACAGTCCGCTCGGCTGCAACGCAATCTTGGTGTTCATGG 496
 DB 4373249 TGAAGCCACCGACAGAGGCTTCGCTCGGCGCAACGCAATCTTGGTGTTCATGG 4373190
 QY 497 CTGTTGCAAGGCTGCTGCTGATTCGCGAGGCTTCCCATCTTCCGCTTACATCGGTGGAC 556
 DB 4373189 CCGTCGCGCAGCGCTTCGAGCGCAGCGACCTCCGCTCTTCCGCTTCTCCGCTTCTCCGCT 4373130
 QY 557 CAACCGCACAGTCTTCCAGTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 616
 DB 4373129 CGAACGCGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4373070
 QY 617 ACTCCGCTGTTGAGTTCAGGAAATTCATGATGATGATGATGATGATGATGATGATGATGATGATG 676
 DB 4373069 ACTCCAGCTGAGATCCAGAGTTTATGATGATGATGATGATGATGATGATGATGATGATGATG 4373010
 QY 677 AGGCTCTCCGCAACGCGCGCGAGGTTTACCAACGCTGAAATCGCTCATCAAGGAAAGG 736
 DB 4373009 AGGCTCTCCGCGCGCGAGGTTTACCAACGCTGAAATCGCTCATCAAGGAAAGG 4372950
 QY 737 GCGTGTCCACCGGACTTGGCGATGAGGCGGCTTCCGCTTCCGCTCGGCTTCCGCTTCCGCTTCCGCT 796
 DB 4372949 GCGTGTCCACCGGCTTGGCGATGAGGCGGCTTCCGCTTCCGCTCGGCTTCCGCTTCCGCTTCCGCT 4372890
 QY 797 AGGCTCTTGAACCTTATCGTTGAGGCAATCCAGAGGCTGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCT 856
 DB 4372889 CCGGCTCGACCTCATCATCGAGGCCATCAAGCAGCGCGTTATATCCCGGCGAGGAGCA 4372830
 QY 857 TCGCTCTTGTCTGAGAGTTCCTTCTGAGTTCCTTCAAGAGCGGCACTTACCATTCG 916
 DB 4372829 TCGGCTCGGCTCGAGCTCGCGCTCGAGTTCCTTCAAGAGCGGCACTTACCATTCG 4372770
 QY 917 AAGTGGCCAGCACTCCGAGCTGAGATGCAACGTTTACGCTGAGCTGTTGACGCTG 976
 DB 4372769 AGGCAAGTCCCGCTCGGCGCGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 4372710
 QY 977 ACCCAATCTGCTCCATCGAGGACCCACTGCAAGAGATGATGATGATGATGATGATGATGATGATGATGATG 1036
 DB 4372709 ACCGCTCGTCTCATCGAGGACCCGCTGTACGAGGAGCTGGGCGGCTGGAGGTCA 4372650
 QY 1037 TCACCGCAACCATCGGCGCAAGGTTTCAAGTTCGTTGCGGAGCTTCTTCTGCTCAACCAACC 1096
 DB 4372649 TCACCGCAACCATCGGCGCAAGGTTTCAAGTTCGTTGCGGAGCTTCTTCTGCTCAACCAACC 4372590
 QY 1097 CTGAGCGCTGAGAGGAGGCTGCTGATGAGGCTGCAACTCCATCTCTGTTAGGTGA 1156

Db 4372589 CGAGGGCTGGCCCGCGGATCGAGGAGGGCTCCGCCAAGCCCTGCTCGTCAAGGTCA 4372530
Qy 1157 ACCAGATCGGTACCTTCAACCGAGACCTTGAAGCTGTGACATGGCTCAACCGCGAGGCT 1216
Db 4372529 ACCAGATCGGTTCGCTGACCGGACCTTGAAGCTGTGACATGGCTCAACCGCGAGGCT 4372470
Qy 1217 ACACCTCCATGATGCCACCGTTCGGGTGAGACCGAGGACACCAATGCTGACCTCG 1276
Db 4372469 TCAAGTGCATGATGTCACCGCTCCGGCGAGACCGAGGACGTCACCATCGCGGACCTCG 4372410
Qy 1277 CAGTTGCACTCAACTGTGTGCGAGATCAAGACTGTGTCTCCAGCAGCTTCCGACCGGTGCG 1336
Db 4372409 CGTGCCTGTAATGCTGCGGTGAGATCAAGACCGGCGCCCGCTCGGACCGGTGCG 4372350
Qy 1337 CAAAGTACAAACAGCTTCTCCGATCGAGCAGTGTCTGGCGAGCGCGCGCTTACGCGAG 1396
Db 4372349 CCAAGTACAAACAGCTGCTGCGATCGAGGAGATCTTCGAGCAGCGCGCGAGTAGCGG 4372290
Qy 1397 GTGCGAGCGCATTCACAGCTTTC 1420
Db 4372289 GCCGCTCGCGCTTCCCGGTTCC 4372266

RESULT 11

US-10-156-761-3520
; Sequence 3520, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 3520
; LENGTH: 1284
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1284)
US-10-156-761-3520

Query Match 44.2%; Score 697; DB 15; Length 1284;
Best Local Similarity 72.3%; Pred. No. 8.3e-204;
Matches 920; Conservative 0; Mismatches 350; Indels 3; Gaps 1;

Qy 151 GTGGTGAATCATGACGATATCGTCCGGAATCTCGACTCCCGCGGTACCCCAACC 210
Db 1 GTGCGCTCCATCGACTCGTCTGTAGCCCGGGAATCTCGACTCCCGAGGCAACCCACG 60
Qy 211 GTGAGGACAGAGTTTCTCTGTATGACGCTTCCACGCTGCGAGGTGTTCCATCCGCG 270
Db 61 GTGAGGTGAGGTGCGCTTCGACGACGCGACGAGGTGTCGCGCGCTCCGTCGCGC 120
Qy 271 GCATCCACCGCGGTCCACGAGCTCATGCTGCGGTGACGCGTG---GGATCGCTACCTG 327
Db 121 GCTCCACCGGTGCTTCGAGGCGCATCGAGCTCCGCGAGCGGTGACCCCAACCGTTACCAG 180
Qy 328 GGCAGGGGTTTGAAGGCACTTGAAACGTCACGAGAAATCGGGCAAGAGCTCGCT 387
Db 181 GGCAGGGGTGTCGAGAAGGCCCTCTCTCGCGCTCATCGAGCAGATCGGCGCGGAGCTCGTC 240

Qy 388 GGCCTAGAGGTGATCGATCAGCGCTCATCGACGAGCAATGATCAAGCTTGAATGCGCAC 447
Db 241 GGCCTAGAGGTGATCGATCAGCGCTCATCGACGAGCAATGATCAAGCTTGAATGCGCAC 300
Qy 448 GGCACAGTCCCGCTGGGTGCAACCGCAATCTTGGTGTTCATCGCTGTTGCAAG 507
Db 301 GACAAAGAGGTCTCGCTCGCGCGCAACCGCAATCTTGGTGTTCATCGCTGTTGCAAG 360
Qy 508 GCTGTGTGATTCGCGAGGCTCCCACTGTTTCGCTACATCGGTGAGCAAAACGACAC 567
Db 361 GCGCTCCGAGGCGAGCAGCTCCCGCTCTTCGCTACCTGGGCGGCGGAAACGCGCAC 420
Qy 568 GTTCTTCAGTTCGAATGATGAACATCATCAACGCTGGCGCTCAGCTGATCTCGGTT 627
Db 421 CTGCTGCGGTTCGATGATGAACATCTCGTGAACGCGCTCCGACGCGGACTCCAACTG 480
Qy 628 GACGTTCAAGGATTCATGATCGCTCCAACTCGGTGAGAGACTTCTCTGAGGCTCTCGC 587
Db 481 GACATCCAGGATTCATGATCGCTCCGATCGGCGGAGTCTTCTTCGAGGCGCTCGGC 540
Qy 688 AACGCGCGGAGGTCTACACGCTCAAGTCCGCTCATCAAGGAAAGGGCTGTCCAC 747
Db 541 TGGGCGCGGAGGTCTACCAACCTCAAGAGGTGCTGAAGACCAAGGGCTGTCCAC 600
Qy 748 GACTTGGCGATGAGGCGGCTTCGCTCTTTCGCTCCAGCTCCACCGTGAAGCTCTTGAC 807
Db 601 GGCCTCGCGAGGAGGCGGCTTCGCGCGAACTCGGAGTCAACCGCGCGCTCGAC 660
Qy 808 CTTATCGTTGAGGCAATCGAGAGGCTTCAACCGGCTTCAACCGGAGGAGGATCGCTTGT 867
Db 661 CTCATCATCGAGGCGCATCAAGCAGGCGGTTTACATCCCGGCGAGCAGATCGCGCTCG 720
Qy 868 CTGAGAGTGTCTCTCTGAGTTCTCAAGGAGCGGCACTTCAACCTCGAAGGTGGCCAG 927
Db 721 CTCGAGTCCGCGCTCGAGTTCTTCAAGGAGCGGAGTACGAGTTCGAGGCGCAAGTCC 780
Qy 928 CACTCCGAGCTGAGATGGCAACGTTTACGCTGAGCTCGTTGACGCTTACCCCAATCGTC 987
Db 781 CGCTCGCGCGCGAGATGACCGAGTACTACGAGGAGTCTCGCTCCGCGTACCCGCTCGTC 840
Qy 988 TCCATCGAGGACCCACTGCGAGGAGATGATCGGAGGTTTACCAACCTTCAACCGCAAC 1047
Db 841 TCCATCGAGGACCCGCTGTACGAGGAGCAGTGGGCGGCTGGAAAGGTTCATCAGCGACA 900
Qy 1048 ATGCGCGACAGGTTTCAAGTGGGCGAGCTTCTTCGTCACCAACCTTGAAGCGCGCTG 1107
Db 901 CTGGGCGACAGGTTCCAGATCGTGGGCGAGCAGCTTTCGTCACCAACCGGAGCGCTG 960
Qy 1108 AAGGAGGCGCATCGCTAAGAGGCTGCGCAACTTCCATCTCTGTTAAGGTGAACAGATCGGT 1167
Db 961 GCCCGCGCATCGAGGAGGCTCCGCGCAACGCGCTTCTCGTCAAGTCAACAGATCGGT 1020
Qy 1168 ACCCTCAACGAGACTTCGAGCTGTGACATGCTCATCGCGGAGGTTACACCTTCCATG 1227
Db 1021 TCGCTGACCGAGACCTTCGAGCGCTCGAGTGGCGCCAGCGCAACGCGTTCAGTGCATG 1080
Qy 1228 ATGTCCACCGTTCGCTGAGACCGAGGACACCACTTCTGCTGACCTCGAGCTTGCACCT 1287
Db 1081 ATGTCCACCGCTTCGCGGAGACCGAGAGCTTCACTTCGCGGACCTCGCGCTCGCGCTG 1140
Qy 1288 AACTGTGGCGAGTCAAGACTGTTGCTCCAGCAGCTTCCGACCGTTCGCAAGGTACAAAC 1347
Db 1141 AACTGTGGTCAATCAAGACCGGCGCGCGCGCTCGGACCGTGTGCGCAAGTACAAAC 1200
Qy 1348 CAGCTTCTCGCATCGAGCAGCTTCTTGGCGAGCGCGCGCTTACGAGGCTCGAGCGCA 1407
Db 1201 CAGCTGTGTGATCGAGGAGATCTTCGAGCAGCGCGCGGAGTACGCGCGCGCTCGCG 1260
Qy 1408 TTCCACGCTTTC 1420
Db 1261 TTCCCGGCTTTC 1273

RESULT 12

US-10-369-493-32027
 ; Sequence 32027, Application US/10369493
 ; Publication No. US20030233675A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Chen, Xianfeng
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
 ; FILE REFERENCE: 38-10(52052)B
 ; CURRENT APPLICATION NUMBER: US/10/369,493
 ; CURRENT FILING DATE: 2003-02-28
 ; PRIOR APPLICATION NUMBER: US 60/360,039
 ; PRIOR FILING DATE: 2002-02-21
 ; NUMBER OF SEQ ID NOS: 47374
 ; SEQ ID NO 32027
 ; LENGTH: 1269
 ; TYPE: DNA
 ; ORGANISM: Thermobifida fusca
 ; US-10-369-493-32027

Query Match 43.1%; Score 679.8; DB 17; Length 1269;
 Best Local Similarity 71.0%; Pred. No. 1.7e-198;
 Matches 900; Conservative 0; Mismatches 367; Indels 0; Gaps 0;

QY	152	TGGCTGAAATCATGACGATTTGCTCGGAAATTCGACCTCCCGCGGTAAACCAACCG	211
DB	2	TGGCTGATCGAGGAGTACACGACGGAGATTCTCGACTCCCGCGGTAAACCAACCG	61
QY	212	TCGAGGACAGAGTTTCTCGATGACGGTTCCACAGGTGTCGACAGGTGTTCCATCCGCG	271
DB	62	TCGAGGTGAGTTCGACTCGACGACGGACGATTGCCCGGCGGTGTCCTCCAGTGGG	121
QY	272	CATCACCGGCTCCAGGAGTCTAGCTCGGTGACGGTGGGATCGCTACCTGGGCA	331
DB	122	CGTCCACCGGTGAGTTCGAAGCGGTGCAACTCGCGGATGTTGGCGACCGTTACGGCGTA	181
QY	332	AGGGGCTTTGAAGCGAGTTGAAGAGTCAAGAGAAATCGGACGAGCTCGCTGCC	391
DB	182	AGGGGCTGAGAGAGCGCTCGCGCGGTCAACAGAGAAATCTCCGAACAGATCTGGGTT	241
QY	392	TAGAGGCTGACGATCAGCGCTTCATCGACGAGCAATGATCAAGCTTGATGACACCGCA	451
DB	242	ACGAGGCGGAGGACGAGCGCTCGTCGACAGCGCTGATCACCTGGACGTTACCCCG	301
QY	452	ACAAGTCCCGCTGGGTGCAACGCAATCTTGGTGTTCATGCTGTTTCAAGAGGTG	511
DB	302	ACAAGTCCCGCTGGGTGCAACGCAATCTTGGTGTTCATGCTGTTTCAAGAGGTG	361
QY	512	CTGCTGATTTCGCGAGCTCCACATGTTCCGCTACATCGGTGGACCAACGACAGCTTC	571
DB	362	CGCGGAGAGCGGAGCTTCGCTGCTTCGCTGCTTCGCTGCTTCGCTGCTTCGCTGCT	421
QY	572	TTCAGTTCCATGATGACATCATCAAGTGGGCTCAGCTGACTCCGGTGTGAGC	631
DB	422	TGCGGCTGCGATGATGAATCTCTCAAGGGGTGCGACGCGGACCAACGAGCTGACGA	481
QY	632	TTGAGGATTCATGATCGCTCCATCGGTGCGAGACCTTCTGAGGCTTCGCGACG	691
DB	482	TCCAGAGTTCATGATCGCTCCATCGGTGCGAGACCTTCTGAGGCTTCGCGACG	541
QY	692	CGCGGAGGTTCACACGACTGAAAGTCCGCTCATCAAGGAAAGGGCTGTCACCGGAC	751
DB	542	CGCGGAGGTTCACACGACTGAAAGTCCGCTCATCAAGGAAAGGGCTGTCACCGGAC	601
QY	752	TGGCGATGAGGCGGCTTCGCTTCCTCGCTGGCTCCACCGGTGAGGCTCTTACCTTA	811
DB	602	TCGGTGAAGGCGGCTTCGCTTCCTCGCTGGCTCCACCGGTGAGGCTCTTACCTTA	661
QY	812	TGTTGAGGCAATCGAAGGCTGCTTACCCCGGAGGACGACATCGCTCTTGTCTGG	871

DB	662	TCAGCGAAGCCATCACGAAGCGCGCTTCACTGGGCCAGGACATCGCCCTGGCTTG	721
QY	872	ACGTTGCTTCTCTGAGTCTTCAAGGACGGCACTTACCACTTCAAGAGTGGCAGCACT	931
DB	722	ACGTTGCGCCACCGAGTCTTACCGGACGGCGCTTACCACTTCAAGAGTGGCAGCACT	781
QY	932	CCGAGCTGAGATGGCAAAAGTTTACGCTGAGTCTGCTTGAACCGCTTACCAATCTCTCCA	991
DB	782	CGCGGAGGAGATGGCGCCTACTACACCGAGCTGCTGAGTCTTACCGCTGCTGCTGA	841
QY	992	TCGAGGACCACTCCAGGAGAGATGCTGGGAGGTTTACACCACTTCAACCGCAACCATCG	1051
DB	842	TCGAGGACCACTCCAGGAGAGATGCTGGGAGGTTTACACCACTTCAACCGCAACCATCG	901
QY	1052	GCACAAAGTTTCAGATCGTTGGGAGGAGCTTCTTCTGTCACCAACCTTGGGAGG	1111
DB	902	GCACAAAGTTTCAGATCGTTGGGAGGAGCTTCTTCTGTCACCAACCTTGGGAGG	961
QY	1112	AGGCATCGCTAAGAGGCTGCCAACTTCCATCTCTGTTAAGGTGAACAGATCGGTACCC	1171
DB	962	GCATATCGAGGAGGCGCTGCCAACTTCTGTTAAGGTGAACAGATCGGTACCC	1021
QY	1172	TCACCGAGCTTCGAGCTGTCGACATGCTCACCGCGAGGCTACACCTTCCATGATGT	1231
DB	1022	TCACCGAGCTTCGAGCTGTCGACATGCTCACCGCGAGGCTACACCTTCCATGATGT	1081
QY	1232	CCACCGTTCGAGCTGAGCGGAGGACCAACCACTTCTGTCGACCACTTGGAGTTCGACT	1291
DB	1082	CCACCGTTCGAGCTGAGCGGAGGACCAACCACTTCTGTCGACCACTTGGAGTTCGACT	1141
QY	1292	GTGCCAGATCAAGACTGCTGCTCCAGCACTTCCGACCGCTGCGAAAGTACAAACGAGC	1351
DB	1142	CGGGGAGATCAAGCGGCTGCTCCGCGCGAGCGAGCGGCTGCGAAAGTACAAACGAGC	1201
QY	1352	TTCTCGCATCGAGCTGCTTGGGAGCGCGGCTTACGAGGCTTACGAGGCTGCGAGGATTC	1411
DB	1202	TGCTCGCATCGAGGAGCTTGGGAGCGCGGCTTACGAGGCTTACGAGGCTGCGAGGATTC	1261
QY	1412	CACGCTT 1418	
DB	1262	CGGTTT 1268	

RESULT 13
 US-10-282-122A-25949
 ; Sequence 25949, Application US/10282122A
 ; Publication No. US20040029129A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Liangsu
 ; APPLICANT: Malone, Cheryl
 ; APPLICANT: Haselbeck, Robert
 ; APPLICANT: Ohlsen, Karl
 ; APPLICANT: Zyskind, Judith
 ; APPLICANT: Wall, Daniel
 ; APPLICANT: Trawick, John
 ; APPLICANT: Carr, Grant
 ; APPLICANT: Yamamoto, Robert
 ; APPLICANT: Forsyth, R.
 ; APPLICANT: Xu, H.
 ; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
 ; FILE REFERENCE: ELITRA.034A
 ; CURRENT APPLICATION NUMBER: US/10/282,122A
 ; CURRENT FILING DATE: 2003-02-20
 ; PRIOR APPLICATION NUMBER: 60/191,078
 ; PRIOR FILING DATE: 2000-03-21
 ; PRIOR APPLICATION NUMBER: 60/206,848
 ; PRIOR FILING DATE: 2000-05-23
 ; PRIOR APPLICATION NUMBER: 60/207,727
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: 60/230,335
 ; PRIOR FILING DATE: 2000-09-06

```

; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 25949
; LENGTH: 1287
; TYPE: DNA
; ORGANISM: Mycobacterium avium
US-10-282-122A-25949

Query Match
Best Local Similarity 41.0%; Score 647; DB 17; Length 1287;
Matches 888; Conservative 0; Mismatches 380; Indels 3; Gaps 1;

QY 151 GTGGCTGAATCATGCACGTATTTCGCTCGCGAAATTCCTCGACTCCCGGGTAAACCCCAACC 210
DB 1 GTGGCGATTATCGACGAGTTCGGGGCCCGCGAGATCCTCGACTCCCGCGGTAAACCCGACA 60

QY 211 GTGAGGCGAGAGGTTTTCCTGGATGACGGTTTCCACGGTGTGCGAGGTTTCCATCCCGGC 270
DB 61 GTGAGGTCGAGATCGCCCTGACCGAGGAAAGTTTCCGCCGCGCGGGTGCCTCGGG 120

QY 271 GCATCACCGCGGTCCACGAGGTATAGCTGCGTGAAGCGTGGGATCGCTACTCTGGGC 330
DB 121 GCGTCACCGGTGAGCAGCGGTGTGCGAACTGCGCAGCGCGGGGAGCGGTACGCGCGC 180

QY 331 AAGGCGTTTTGAAGCGAGTTTGAACGTTCAACGAGAAATCGGACGAGTTCGCTGCG 390
DB 181 AAGGCGTGCAGAGCGGTGCGGCGGTGCGGCGGTGCGGAGATCGGCGCGCGGTGATCGGG 240

QY 391 CTAGAGGTGACGATCAGCGCTCTACGACGAGCAATGATCAAGCTTGATGGCACCGCC 450
DB 241 CTGACGCGCAGCAGCAGCGCTGTGTGACGAGCGCTGCTGACCTGACGCGCACCGCG 300

QY 451 AACAACTCCCGCTGGGTGCAACGCAATCTTGGTGTGTTTCCATGCTGTGTTGCAAGGCT 510
DB 301 GACAAAGTCGAGCTGGCGCGCAACGCAATCTTGGGCGTGTGCTGGCGGTGCGCAAGCG 360

QY 511 GCTGCTGATTCGCGAGGCTCCACTGTTCCGCTATCATGCTGGTGGAGCAACGCAAGCT 570
DB 361 GCGCGCGATTCGCGGAGTGTGCGGTGTTCCGCTACCTGGGCGGCGCCACGCGCACATC 420

QY 571 CTTCCAGTTCCTCAATGATGAATCATCAACGCTGGCGCTCAGCTGACTCCGGTGTGAC 630
DB 421 CTGCGCGTCCGATGATGAATCTTCAACGCGGTGCGCAGCGCACACCGCGCTGCAC 480

QY 631 GTTCCAGGAATTCATGATCGCTCCATTCGCTGCGAGACCTTCTCTGAGGCTCTCCGCAAC 690
DB 481 ATCCAGGAGTTCATGTCGCCCCCGATCGCGCGCGCGAGTTTCGCGGAGGATTCGCTGG 540

QY 691 GCGCGGAGGTCTACGACGATGAACTGAGTCCGTCATCAGGAAAGGCTGTGTCACCGGA 750
DB 541 GGTGCGGAGGTGATCACTCGTGAAGTGTGCTGCTGAAGAGGAGGCGCTGAGCACCGCGC 600

QY 751 CTTGCGGATGAGGCGGCTTCCCTCTTCCGCTCCACCGGTCCACCGGTGAGGCTTTGACCT 810
DB 601 CTGCGCGAGAGGCGGCTTCCCGCGGACGTGGCGGGCACACCGCGCGCTGAGCACTG 660

QY 811 ATCGTTGAGGCAATCGAAGAGCTGGCTTCAACCGCGGAGGACATCGCTCTTCTGCTG 870
DB 661 ATCGCGCGGCGCATCGAATCGGCGCGGCTTCAACTCGGCGACCGACGATGCGGCTGCGCTC 720

871 GACGTTGCTCTCTGAGTTCTTCAAGGACGCGCAC---CTACCATTCTGAAAGGTGGCCAG 927
721 GACGCGGGGGCCACCGAGTTTCTACAGCGACCGGCTACAGTTTCGAGGGGAGCACC 780
928 CACTCCGAGCTGAGATCGCAAAACGTTTACGCTGAGCTCGTTGACGCGTACCCTCAATCGTC 987
781 CGCACCGCGGAGCAGATGCGCGAGTTTCTACGCGGGCTGCTCGCGCGCTATCCGTTGGTG 840
988 TCCATCGAGGACCCACTCGCAGGAGATGACTGCGGAGGTTTACCAACCTCACCGCAACC 1047
841 TCCATCGAAGACCCCGCTGTCCGAAGATGATTGGGACGCGCTGGGCGGCGCTGACCGCGTCG 900
1048 ATCGCGCAAAAGTTTCAGATCGTTGGCGACGACTTCTTCTGTCACCAACCTCGAGCGCTG 1107
901 ATCGCGGACCGGGTGCAGCTGTGCGCGACGACGCTCTTCTGTCACGAAACCCCGAAGCGCTG 960
1108 AAGGAGGCGATCGCTTAAGAAGCTGCGCAACTCCATCTCTGGTTAAGGTGAACCCAGATCGGT 1167
961 GAAGAGGCTATCGAGAAGGGCGTGCAGAAATCGTTGCTGCTCAAGGTGAATCAGATCGCG 1020
1168 ACCCTCAGGAGACTTCGACGCTGTGACATGCTCACCGGCGAGGCTACACCTCCATG 1227
1021 ACCTGACGAGACGCTGGACGCGCTCGCTGCGGCCACACAGCGGCTACCGCACGATG 1080
1228 ATGTCCCAACCGTTCCGCTGAGACGAGGACACCACTTGTGCTGACCTCGAGTTGCACCTC 1287
1081 ATGAGCCACCGAGCGGTGAACCGAGGACACAGATCGCGGACCTGCGGTCGCGGTC 1140
1288 AACTGTGCGCAGATCAAGACTGTGCTCCAGCAGCTTCCGACCGCTGTCGCAAGGTACAAAC 1347
1141 GCGAGCGGCGAGATCAAGACCGGAGCGCGCGCGCGAGCGGCTGCGCAGGTACAAAC 1200
1348 CAGCTTCTCCGATCGAGCAGCTGTGCGGACGCGCGCGCTTACGAGGTGCGAGCGCA 1407
1201 CAGCTGTGCGGATCGAGGAGCGCTCGGCGACGCGCGCTTACGCGGCGGACCTGGCC 1260
1408 TTCCCGCGCTT 1418
1261 TTCCCGCGCTT 1271

RESULT 14
US-10-080-170-649
; Sequence 649, Application US/10080170
; Publication No. US20030129601A1
; GENERAL INFORMATION:
; APPLICANT: COLE, S.T.
; TITLE OF INVENTION: COMPARATIVE MYCOBACTERIAL GENOMICS AS A TOOL FOR
; IDENTIFYING TARGETS FOR THE DIAGNOSIS, PROPHYLAXIS OR
; TREATMENT OF MYCOBACTERIOSES
; FILE REFERENCE: 03495.0218
; CURRENT APPLICATION NUMBER: US/10/080,170
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: 60/270,123
; PRIOR FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 652
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 649
; LENGTH: 45191
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
US-10-080-170-649

Query Match
Best Local Similarity 40.1%; Score 632.6; DB 15; Length 45191;
Matches 891; Conservative 0; Mismatches 409; Indels 3; Gaps 1;

QY 140 AGGAGGCGCAGTGGCTGAAATCATGACGATTCGCTGCGGAAATTCGACTCCCGCG 199
DB 19931 AGGAGAACCCAGTCCCGATTATCGAGCAGGTTAGGCGCGGAGAGATCTCTGATTCGCG 19990
200 GTAACCAACCGTCGAGCAGAGGTTTCTGATGACGTTTCCACGCGGTGCGAGGTG 259
```

Db 19991 GCAACCCGACGCTGGAGTTCGAGTGGCGCTTATCGACGGGACATTCGCCCGGCGCGG 20050
Qy 260 TTCCATCCGCGCATCCACCGCGCTCCACGAGGCTCATGAGCTGCGTGAACGGTGGCGATC 319
Db 20051 TCCCGTCCGCGCGCTCGACCGGGAGCACAGCGCGTCCAGTTCGCGCACGCGCGCGATC 20110
Qy 320 GCTACTCCGCGCAAGCGCTTTTGAAGGCGAGTTGAAACGTCACAAAGAAATCCGCGACG 379
Db 20111 GCTACGGCGCAAGCGCTGCAAAAGCGCTGACGGCTGTTCTTGATGAGATCGCGCGG 20170
Qy 380 AGCTCGCTGGCTAGAGGCTGACGATCAGCGCTCATCGACGAAGCAATGATCAAGCTTG 439
Db 20171 CGGTCAATCGACTCAACGCGACGACGAGCGATTGGTCGACAGCGCTGGTGGACCTAG 20230
Qy 440 ATGGACACCGCAACAGTCCCGCTGGGTGCAACGCAATCTTGGTGTTCATGGCTG 499
Db 20231 ACGGCACCCCGACAAGTCCCGCTGGGCGGCAACGCGATCTTGGGTGTCTCGCTCGCTG 20290
Qy 500 TTGCAAGGCTGCTGCTGATTCGCGAGCGCTCCCACTGTTCCGCTACATCGTGGTGAACAA 559
Db 20291 TTGCCAAGCGCGCGCGGATTCGGGCGAGCTCCCGTGTTCGCTTATGTCCGGGGGCCAA 20350
Qy 560 AGGCACACGTTTCTTCAGTTCCATGATGAACATCATCAACGGTGGCGCTCACGCTGACT 619
Db 20351 ACGGCACATCTGCGCGTACCGATGATGAACATCTCAACGGCGGCGCACACGCCGATA 20410
Qy 620 CGGTGTTGAGTTGAGGAATTCATGATCGCTCCATCGGTGCGAGACCTTCTCTGAGG 679
Db 20411 CGGTGTGCAATTCAGAGTTTCATGTTGGCGGCAATTTGGCGCGCCACGCTTCTCGAGG 20470
Qy 680 CTCTCCGCAACGGCGCGAGGTCTACACGACACTGAGTCCGTATCAAGGAAAGGGCC 739
Db 20471 CGTTGGCTGGGCGCTGAGTGTATCAACGCGCTCAAGTCGCTGTAAGAGAGGGGC 20530
Qy 740 TGTCCACCGGACTTGGCGATGAGGCGGCTTTCGCTCTTCGCTCGGCTCCACCGTGAAG 799
Db 20531 TGTCCACCGGCTTGGCGACGAAGGCGCTTTCGCGCGGATGTGCGCGCACCAACCGCG 20590
Qy 800 CTCTTGACCTTATCGTTGAGCGAATCGAAGGCTGGCTTCAACCGAGGCAAGACATCG 859
Db 20591 CGTTGGACCTGATCAGCGCGGCGATCGAGTTCGGCGGCTTTCGACCGCGCGCGACGTG 20650
Qy 860 CTCTTGCTCTGAGCTTCTCTCTGAGTTCTTCAAGGAGCGGAC ---CTACACCTTCG 916
Db 20651 CGCTGGCGCTGAGCGCGCGCGCACCGAGTTCTTCAACGACGCGACCGCTACGCTTCG 20710
Qy 917 AAGTGGCGACGACTCCGACGCTGAGATGGCAACGTTTACGCTGAGCTCGTTGACGCGT 976
Db 20711 AGGGCACCCCGTACCGACACGATGACCGAGTTCTACGCGGCGCTGCTCGCGCGCT 20770
Qy 977 ACCCAATCGTCTCCATCGAGGACCCACTGACAGGAAGATGATGGAGGCGTTTACACCAAC 1036
Db 20771 ACCCGCTGGTGCATCGAAGACCCACTGTTCGAGAGACGATGGAGCGGCTGGCGCGCG 20830
Qy 1037 TCACCGCAACCATCGGCGCAAGGTTGAGTCTGTTGCGAGCTTCTTGTTCACCAAC 1096
Db 20831 TGAACGCGCTCGATCGGTGACCGGCTGCAATCTGCGGCGACGACATTTTGTCAACCAATC 20890
Qy 1097 CTGAGCGCTGAAGAGGCGCATCGCTAAGAGGCTGCAACTCCATCCTGTTAAGGTGA 1156
Db 20891 CCGAGCGGCTCGAGAGGCGCATCGAAGCGGCGCTGGCAATTCGCTGCTGCTCAAGGTGA 20950
Qy 1157 ACCAGATCGGTACCTTCACGAGACCTTCGACGCTGTCGATGGCTCAGCGGCGAGCT 1216
Db 20951 ACCAGATCGGAGCTTGAACGAGACACTCGACGCGGTCACGCTGGCTCAACACGGGGAT 21010
Qy 1217 ACACCTCCATGATTCACCGCTTCGCGTGAACCGGAGACACCACTTGTCTGACCTCG 1276
Db 21011 ACCGACGATGATCAGTCAACCGGAGTGGCGAGACGAGGACACCATGATCGCGGACCTCG 21070
Qy 1277 CAGTTGACCTCACTGTGGCGGAGATCAAGACTGTGTCTCAGACAGCTTCCGACGCTGTCG 1336
Db 21071 CGGTGGCCATCGGACGCGGCGAGATCAAGACGCGCGCGCTGCTCGCAGTGAAGCGCGCTCG 21130

Qy 1337 CAAAGTACAAACGAGTTCCTCCGATCGAGCAGCTGCTTGGCGACGCGCGCTTACGCGAG 1396
Db 21131 CAAATACAAACGAGTTCCTCCGATCGAGCAGCTTGGCGACGCGCGCTTACGCGG 21190
Qy 1397 GTCCGACGCGATTCACGAGCTTTCAGGGCTTAATAAAGCGCT 1439
Db 21191 GCGACCTGGCAATTCCTCCGTTCCGTCGCGAGACGAATAGGT 21233

RESULT 15

US-10-080-170-649

; Sequence 649, Application US/10080170

; Publication No. US20040121322A9

; GENERAL INFORMATION:

; APPLICANT: COLE, S.T.

; TITLE OF INVENTION: COMPARATIVE MYCOBACTERIAL GENOMICS AS A TOOL FOR

; TITLE OF INVENTION: IDENTIFYING TARGETS FOR THE DIAGNOSIS, PROPHYLAXIS OR

; TITLE OF INVENTION: TREATMENT OF MYCOBACTERIOSES

; FILE REFERENCE: 03495.0218

; CURRENT APPLICATION NUMBER: US/10/080,170

; PRIOR FILING DATE: 2002-06-10

; PRIOR APPLICATION NUMBER: 60/270,123

; PRIOR FILING DATE: 2001-02-22

; NUMBER OF SEQ ID NOS: 652

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 649

; LENGTH: 45191

; TYPE: DNA

; ORGANISM: Mycobacterium tuberculosis

US-10-080-170-649

Query Match

Best Local Similarity 40.1%; Score 632.6; DB 19; Length 45191;

Matches 891; Conservative 0; Mismatches 409; Indels 3; Gaps 1;

Qy 140 AGGAGGCCACAGTGGCTGAATCATGCACGTATTTCGCTCGGAAATTCGACTCCGCG 199
Db 19931 AGGAGAACCCAGTGGCGGCTTATCGAGCAGTATTAGGCGCCGAGAGATCTCGATTCCCGCG 19990
Qy 200 GTAACCCAAACCGTCGAGGCGAGAGGTTTCTCGATGACGCTTCCACGCTGTCGACAGTG 259
Db 19991 GCNACCCGACGCTGAGGCTCGAGGTCGAGGTCGCGCTTATCGACGGGACATTCGCCCGCGCG 20050
Qy 260 TTCCATCCGCGCATCCACCGCGCTCCACGAGCTCATGAGCTGCGTGAACGCTGCGGATC 319
Db 20051 TGCCTGCGCGCGCTCGACCGGCGACGAGCGCTGAGTTGCGCGACGCGCGGATC 20110
Qy 320 GCTACTGGCGCAAGGCGCTTTTGAAGGCGAGTTGAAAAAGTCAACGAGAAATCGCGGACG 379
Db 20111 GCTACGGCGCAAGGCGCTGCAAAAAGCGCTGCAAGCTTCTTGTATGATGATCGGCGCG 20170
Qy 380 AGCTCGCTGGCTTAGAGGCTGACGATCAGCGCTCATCGACGAAGCAATGATCAAGCTTG 439
Db 20171 CGGTCAATCGACTCAACGCGGACGACGACGAGTGGTCGACCGGCGCTGGTGGACCTAG 20230
Qy 440 ATGGCACCGCCAAAGTCCCGCTGGGTGCAACGCAATCTTGTGTTCATGGCTG 499
Db 20231 ACGGCACCCCGACAAAGTCCCGCTGGGCGGCAACGCGATCTTGGGTGCTCGCTCGCTG 20290
Qy 500 TTGCAAGGCTGCTGCTGATTCGCGAGGCTCCCACTGTTCCGCTACATCGGTGGACCAA 559
Db 20291 TTGCCAAGCGCGCGGAGTTCGGGCGAGCTCCCGTGTTCGCTTATGTCCGGGGGCCAA 20350
Qy 560 ACGCACACGTTTCTTCAGTTTCCAATGATGAACATCATCAACGGTGGCGCTCACGCTGACT 619
Db 20351 ACGGCACATCTGCGCGTACCGATGATGAACATCTCTCAACGCGCGGCGCACACGCCGATA 20410
Qy 620 CGGTGTTGACCTTCAGGAATTCATGATCGCTCCATTCGATCGGTGCGAGACCTTCTCTGAGG 679
Db 20411 CGGTGTGACATTCAGAGTTTCATGGTGGCGCAATTTGGCGCGCCACGCTTCTGTCGAGG 20470
Qy 680 CTCTCCGCAACGGCGCGAGGTCTACACGCACTGAAAGTCCGCTCATCAAGGAAAGGGCC 739

Db	20471	CGTTGGCTGGGGCGCTGAGGTGTACACGCGCTCAAGTCGGTCTGAAAAAGGAGGGGC	20530
Qy	740	TGTCCACCGGACTTGGCGGATGAGGCGGCTTGGCTCTTCCTCGGCTCCACCCGTGAGG	799
Db	20531	TGTCCACCGGCTTGGCGGATGAGGCGGCTTGGCTCTTCCTCGGCTCCACCCGTGAGG	20590
Qy	800	CTCTTGACCTTATCGTTGAGGCAATCGAAGAGGCTGGCTTACCCGAGGCAAGGACATCG	859
Db	20591	CGTTGGACCTGATCAGCCGGGCTTCGAGTGGCGGGCTTGGACCCGCGCGGACGCTGG	20650
Qy	860	CTCTTGCTCTGAGCGTGTCTCTGAGTTCTTCAAGGACGGCAC---CTACCACTTCG	916
Db	20651	CGCTGGCCCTGACGCGCGGCGCACCGAGTTCTTACCGACGGCACCGGCTACGCTTCG	20710
Qy	917	AGGTGGCCAGCACTCCGCGAGCTGAGTGGCAACGTTTACGCTGAGCTCGTTGACGGT	976
Db	20711	AGGGCACCAACCCGTACCGGACGACCGAGTTCTACCGGGCTTGTCCGGCGCT	20770
Qy	977	ACCCAATCGTCTCCATCGAGGACCCACTGCAGGAGATGACTGGGAGGTTACACCAACC	1036
Db	20771	ACCCGCTGGTGTGATCGAGACCCACTGTCCGAGACGATTTGGACGSGCTGGGCGGCG	20830
Qy	1037	TCACCGCAACCATCGCGCAAGGTTACAGTCTTGGCGACGACTTCTTCGTCAACCAACC	1096
Db	20831	TGACGGCTCGATCGGTGACCGGTCGAAATCGTGGCGACGACATCTTTGTCAACAATC	20890
Qy	1097	CTGAGGGCTGAAGAGGGCATCGCTAAGAGGCTGCCAACTCCATCTCTGGTTAAGGTGA	1156
Db	20891	CCGAGCGGCTCGAGGAGGSCATCGAAACGGGGGCTGGCAAAATCGTTGTGTCAAGGTGA	20950
Qy	1157	ACCAGATCGGTACCTCACCGAGACCTTCGACGCTGTGACATGGCTACCGCGCAGGCT	1216
Db	20951	ACCAGATCGGACGTTGACCGGACACTCGACGCGTACGCTGGCTACCAACGGCGGAT	21010
Qy	1217	ACACCTCCATGATGTCCACCGTTCCGGTGAGACCGAGGACACCACTTGTGACCTCG	1276
Db	21011	ACCGCACGATGATCAGTCACCGCAGTGGCGGAGACGAGGACACCATGATCGCGGACCTCG	21070
Qy	1277	CAGTTGCACTCAACTGTGGCCAGATCAAGACTGGTGCTCCAGCAGTTCCGACCGGTGCG	1336
Db	21071	CGGTGGCCATCGGACGCGGCGAGATCAAGACGGGGCGGCTCTGCTGAGTGAAGCGGTCG	21130
Qy	1337	CAAGGTACAACCAAGCTTCTCCGATCGAGCAGCTGTTGGCGACCGCGGCTTACGCGAG	1396
Db	21131	CAAAATACAACCAAGCTGTTGGGATCGAAGAGGCGCTTGGCGACCGCGGCTTACGCGG	21190
Qy	1397	GTGCGAGGCGATTCCCAAGCTTTCAGGGCTAAATAAAGCGCT	1439
Db	21191	GGGACCTGGCATTTCTCGGTTCCGCTGCGAGACGAAATAGGT	21233

Search completed: September 29, 2005, 11:59:52
Job time : 4260.5 secs

THIS PAGE BLANK (user)

THIS PAGE BLANK (user)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 29, 2005, 12:13:58 ; Search time 1026 Seconds
(without alignments)
171.954 Million cell updates/sec

Title: US-10-728-947-2
Perfect score: 2155
Sequence: 1 VAEIMHVAREILDSRGNPT.....QLLGAGVYAGRSAPPRFQG 425

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1846076 seqs, 415116000 residues

Total number of hits satisfying chosen parameters: 1846076

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_AA:
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US10F_PUBCOMB.pep.*
19: /cgn2_6/ptodata/2/pubpaa/US11A_PUBCOMB.pep.*
20: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
21: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
22: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No: is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2155	100.0	425	9 US-09-860-768-2	Sequence 2, Appli
2	2155	100.0	425	16 US-10-728-947-2	Sequence 2, Appli
3	2152	99.9	425	9 US-09-738-626-4585	Sequence 4585, Ap
4	2151	99.8	425	9 US-09-860-768-4	Sequence 4, Appli
5	2151	99.8	425	16 US-10-728-947-4	Sequence 4, Appli
6	2149	99.7	425	16 US-10-781-014-72	Sequence 72, Appli
7	2149	99.7	425	17 US-10-494-836-8	Sequence 8, Appli
8	1917	89.0	425	15 US-10-282-122A-53908	Sequence 53908, A
9	1566	72.7	423	15 US-10-369-493-8340	Sequence 8340, Ap
10	1558.5	72.3	429	15 US-10-282-122A-62133	Sequence 62133, A
11	1555.5	72.2	428	14 US-10-156-761-11070	Sequence 11070, A

ALIGNMENTS

RESULT 1
US-09-860-768-2
; Sequence 2, Application US/09860768
; Patent No. US20020082403A1
; GENERAL INFORMATION:
; APPLICANT: Mckel, Bettina
; APPLICANT: Pfefferle, Walter
; APPLICANT: Hermann, Thomas
; APPLICANT: Pohler, Alfred
; APPLICANT: Kalinowski, Jorn
; APPLICANT: Bathe, Brigitte
; TITLE OF INVENTION: New Nucleotide Sequences that Code for the Eno Gene
; FILE REFERENCE: 21123/278404
; CURRENT APPLICATION NUMBER: US/09/860,768
; CURRENT FILING DATE: 2001-05-21
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 425
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-860-768-2

Query Match 100.0%; Score 2155; DB 9; Length 425;
Best Local Similarity 100.0%; Pred. No. 2.6e-171;
Matches 425; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 VAEIMHVAREILDSRGNPTVEAEVFLDDGSHGVAGVPSGASTGVHAEHLDDGGDRLG 60
Db 1 VAEIMHVAREILDSRGNPTVEAEVFLDDGSHGVAGVPSGASTGVHAEHLDDGGDRLG 60
Qy 61 KGVLKAVENVNEEIGDELAGEADDQRLIDEAMIKLDTANKSRIGANAILGVSNVAVAKA 120
Db 61 KGVLKAVENVNEEIGDELAGEADDQRLIDEAMIKLDTANKSRIGANAILGVSNVAVAKA 120

Sequence 64490, A
Sequence 62696, A
Sequence 63882, A
Sequence 11896, A
Sequence 16911, A
Sequence 17433, A
Sequence 10557, A
Sequence 17886, A
Sequence 11819, A
Sequence 52760, A
Sequence 14739, A
Sequence 15222, A
Sequence 23299, A
Sequence 46167, A
Sequence 42225, A
Sequence 21006, A
Sequence 19491, A
Sequence 60954, A
Sequence 51920, A
Sequence 47804, A
Sequence 9097, Ap
Sequence 4688, Ap
Sequence 7447, Ap
Sequence 49282, A
Sequence 38, Appl
Sequence 5012, Ap
Sequence 10834, A
Sequence 42554, A
Sequence 7762, Ap
Sequence 9970, Ap
Sequence 32, Appl
Sequence 20662, A
Sequence 23433, A

12 1551.5 72.0 15 US-10-282-122A-64490
13 1551.5 72.0 15 US-10-282-122A-62696
14 1501.5 69.7 15 US-10-282-122A-63882
15 1418 65.8 15 US-10-369-493-11696
16 1408.5 65.4 15 US-10-369-493-16911
17 1406 65.2 15 US-10-369-493-17433
18 1402.5 65.1 15 US-10-369-493-10557
19 1402.5 65.1 15 US-10-369-493-17886
20 1400 65.0 15 US-10-369-493-11819
21 1396.5 64.8 15 US-10-282-122A-52760
22 1390 64.5 15 US-10-369-493-14739
23 1390 64.5 15 US-10-369-493-15222
24 1381.5 64.1 15 US-10-369-493-23299
25 1377 63.9 15 US-10-282-122A-46167
26 1370.5 63.6 15 US-10-282-122A-52225
27 1366.5 63.4 15 US-10-369-493-21006
28 1361.5 63.2 15 US-10-369-493-19491
29 1360.5 63.1 15 US-10-282-122A-60954
30 1358.5 63.0 15 US-10-369-493-21730
31 1358.5 63.0 15 US-10-282-122A-51920
32 1352.5 62.8 15 US-10-282-122A-47804
33 1351.5 62.7 15 US-10-369-493-9097
34 1348.5 62.6 15 US-10-369-493-4688
35 1348.5 62.6 15 US-10-369-493-7447
36 1348.5 62.6 15 US-10-282-122A-49282
37 1346.5 62.5 15 US-10-375-266-38
38 1346 62.5 15 US-09-815-242-5012
39 1346 62.5 15 US-09-815-242-10854
40 1346 62.5 15 US-10-282-122A-42554
41 1345.5 62.4 15 US-10-369-493-7762
42 1336.5 62.0 15 US-10-369-493-9970
43 1336.5 62.0 15 US-10-369-493-32
44 1327.5 61.6 15 US-10-369-493-20662
45 1326 61.5 15 US-10-369-493-23433

```

Qy 121 AADGAGLPLFRYIGGPNHVLVPMNNIINGGAHDSGVGVQEFPIAGETPSALRN 180
Db 121 AADGAGLPLFRYIGGPNHVLVPMNNIINGGAHDSGVGVQEFPIAGETPSALRN 180
Qy 181 GAETVHALKSVIKKGLSTGLDGGFAPSVGSTREALDLIVEAEKAGFTPGKDIALAL 240
Db 181 GAETVHALKSVIKKGLSTGLDGGFAPSVGSTREALDLIVEAEKAGFTPGKDIALAL 240
Qy 241 DVASSEFFKDGTYHFEGGQSAEMANVYAEVLVDAYPIVSIEDPLQEDDWEGYTNLTATI 300
Db 241 DVASSEFFKDGTYHFEGGQSAEMANVYAEVLVDAYPIVSIEDPLQEDDWEGYTNLTATI 300
Qy 301 GDKQVIVGDDFFVTNPERLKEGIAKKAANSILVKVNIQIGTLTETFDVDMHAGYTSMM 360
Db 301 GDKQVIVGDDFFVTNPERLKEGIAKKAANSILVKVNIQIGTLTETFDVDMHAGYTSMM 360
Qy 361 SHRGTEDETTIADLAVALNCGQIKTGAPARSDRVAKYNQILRIEQLLGDAGVYAGRSAP 420
Db 361 SHRGTEDETTIADLAVALNCGQIKTGAPARSDRVAKYNQILRIEQLLGDAGVYAGRSAP 420
Qy 421 PRFOG 425
Db 421 PRFOG 425

RESULT 2
US-10-728-947-2
; Sequence 2, Application US/10728947
; Publication No. US2004020394A1
; GENERAL INFORMATION:
; APPLICANT: Mockett, Bettina
; APPLICANT: Pfefferle, Walter
; APPLICANT: Hermann, Thomas
; APPLICANT: Pohler, Alfred
; APPLICANT: Kalinowski, Jörn
; APPLICANT: Bathe, Brigitte
; TITLE OF INVENTION: New Nucleotide Sequences that Code for the Eno Gene
; FILE REFERENCE: 21123/278404
; CURRENT APPLICATION NUMBER: US/10/728,947
; CURRENT FILING DATE: 2003-12-08
; PRIOR APPLICATION NUMBER: US/09/860,768
; PRIOR FILING DATE: 2001-05-21
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 425
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-10-728-947-2

Query Match 100.0%; Score 2155; DB 16; Length 425;
Best Local Similarity 100.0%; Pred. No. 2.6e-171;
Matches 425; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VAEIMHVFAREILDSRGNTPTVEAEVFLDDGSHGVAGVPSGASTGVHEAHLRDGGDRYLG 60
Db 1 VAEIMHVFAREILDSRGNTPTVEAEVFLDDGSHGVAGVPSGASTGVHEAHLRDGGDRYLG 60
Qy 61 KGVLKAVENVNEIIGDELGLADDDRLIDEAMIKLDGTANKSRIGANAILGVSMVAKA 120
Db 61 KGVLKAVENVNEIIGDELGLADDDRLIDEAMIKLDGTANKSRIGANAILGVSMVAKA 120
Qy 121 AADGAGLPLFRYIGGPNHVLVPMNNIINGGAHDSGVGVQEFPIAGETPSALRN 180
Db 121 AADGAGLPLFRYIGGPNHVLVPMNNIINGGAHDSGVGVQEFPIAGETPSALRN 180
Qy 181 GAETVHALKSVIKKGLSTGLDGGFAPSVGSTREALDLIVEAEKAGFTPGKDIALAL 240
Db 181 GAETVHALKSVIKKGLSTGLDGGFAPSVGSTREALDLIVEAEKAGFTPGKDIALAL 240
Qy 241 DVASSEFFKDGTYHFEGGQSAEMANVYAEVLVDAYPIVSIEDPLQEDDWEGYTNLTATI 300
Db 241 DVASSEFFKDGTYHFEGGQSAEMANVYAEVLVDAYPIVSIEDPLQEDDWEGYTNLTATI 300

```

```

Db 241 DVASSEFFKDGTYHFEGGQSAEMANVYAEVLVDAYPIVSIEDPLQEDDWEGYTNLTATI 300
Qy 301 GDKQVIVGDDFFVTNPERLKEGIAKKAANSILVKVNIQIGTLTETFDVDMHAGYTSMM 360
Db 301 GDKQVIVGDDFFVTNPERLKEGIAKKAANSILVKVNIQIGTLTETFDVDMHAGYTSMM 360
Qy 361 SHRGTEDETTIADLAVALNCGQIKTGAPARSDRVAKYNQILRIEQLLGDAGVYAGRSAP 420
Db 361 SHRGTEDETTIADLAVALNCGQIKTGAPARSDRVAKYNQILRIEQLLGDAGVYAGRSAP 420
Qy 421 PRFOG 425
Db 421 PRFOG 425

RESULT 3
US-09-738-626-4585
; Sequence 4585, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIALI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 4585
; LENGTH: 425
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-4585

Query Match 99.9%; Score 2152; DB 9; Length 425;
Best Local Similarity 99.8%; Pred. No. 4.7e-171;
Matches 424; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VAEIMHVFAREILDSRGNTPTVEAEVFLDDGSHGVAGVPSGASTGVHEAHLRDGGDRYLG 60
Db 1 VAEIMHVFAREILDSRGNTPTVEAEVFLDDGSHGVAGVPSGASTGVHEAHLRDGGDRYLG 60
Qy 61 KGVLKAVENVNEIIGDELGLADDDRLIDEAMIKLDGTANKSRIGANAILGVSMVAKA 120
Db 61 KGVLKAVENVNEIIGDELGLADDDRLIDEAMIKLDGTANKSRIGANAILGVSMVAKA 120
Qy 121 AADGAGLPLFRYIGGPNHVLVPMNNIINGGAHDSGVGVQEFPIAGETPSALRN 180
Db 121 AADGAGLPLFRYIGGPNHVLVPMNNIINGGAHDSGVGVQEFPIAGETPSALRN 180
Qy 181 GAETVHALKSVIKKGLSTGLDGGFAPSVGSTREALDLIVEAEKAGFTPGKDIALAL 240
Db 181 GAETVHALKSVIKKGLSTGLDGGFAPSVGSTREALDLIVEAEKAGFTPGKDIALAL 240
Qy 241 DVASSEFFKDGTYHFEGGQSAEMANVYAEVLVDAYPIVSIEDPLQEDDWEGYTNLTATI 300
Db 241 DVASSEFFKDGTYHFEGGQSAEMANVYAEVLVDAYPIVSIEDPLQEDDWEGYTNLTATI 300
Qy 301 GDKQVIVGDDFFVTNPERLKEGIAKKAANSILVKVNIQIGTLTETFDVDMHAGYTSMM 360

```

Db 301 GDKVQIVGDDFFVTNPERLKEGIAKKAANSILVKVNIQIGTLTETFDVDMHRAAGYTSNM 360
 QY 361 SHRSGETEDTTIADLAVALNCGQIKTGAPASDRVAKYNOLLRIEQLLDAGVYAGRSAP 420
 Db 361 SHRSGETEDTTIADLAVALNCGQIKTGAPASDRVAKYNOLLRIEQLLDAGVYAGRSAP 420
 QY 421 PRFQ 425
 Db 421 PRFQ 425

RESULT 4
 US-09-860-768-4
 ; Sequence 4, Application US/09860768
 ; Patent No. US20020082403A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Mockel, Bettina
 ; APPLICANT: Pfefferle, Walter
 ; APPLICANT: Hermann, Thomas
 ; APPLICANT: Pöhler, Alfred
 ; APPLICANT: Kalinowski, Jorn
 ; APPLICANT: Bathe, Brigitte
 ; TITLE OF INVENTION: New Nucleotide Sequences that Code for the Eno Gene
 ; FILE REFERENCE: 21123/278404
 ; CURRENT APPLICATION NUMBER: US/09/860,768
 ; CURRENT FILING DATE: 2001-05-21
 ; NUMBER OF SEQ ID NOS: 6
 ; SOFTWARE: Patent in version 3.0
 ; SEQ ID NO 4
 ; LENGTH: 425
 ; TYPE: PRT
 ; ORGANISM: Corynebacterium glutamicum
 US-09-860-768-4

Query Match 99.8%; Score 2151; DB 9; Length 425;
 Best Local Similarity 99.8%; Pred. No. 5.6e-171;
 Matches 424; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VAEIMHVFAREILDSRGNTVFAEVLDDGSHGVAGVPSGASTGVHEAHELDDGDRYL 60
 Db 1 VAEIMHVFAREILDSRGNTVFAEVLDDGSHGVAGVPSGASTGVHEAHELDDGDRYL 60
 QY 61 KGVLKAVENNEIEIGDELAGLEADDDORLIDEAMIKLDGTANKSRILGANAILGVSMVAKA 120
 Db 61 KGVLKAVENNEIEIGDELAGLEADDDORLIDEAMIKLDGTANKSRILGANAILGVSMVAKA 120
 QY 121 AADSAGLPLFRYIGGPNHVLVPPMNIINGGAHDSGVVDVQEFMIAPITCAETTFSEALRN 180
 Db 121 AADSAGLPLFRYIGGPNHVLVPPMNIINGGAHDSGVVDVQEFMIAPITCAETTFSEALRN 180
 QY 181 GAEVTHALKSVIKKGLSTGLDGGFAPSVGSTREALDLIVEATEKAGTTPGKDIALAL 240
 Db 181 GAEVTHALKSVIKKGLSTGLDGGFAPSVGSTREALDLIVEATEKAGTTPGKDIALAL 240
 QY 241 DVASSEFFKDGTYHFEQGHSAEAMANYAELVDAYPIVSIEDPLQEDDWEQYTNLTATI 300
 Db 241 DVASSEFFKDGTYHFEQGHSAEAMANYAELVDAYPIVSIEDPLQEDDWEQYTNLTATI 300
 QY 301 GDKVQIVGDDFFVTNPERLKEGIAKKAANSILVKVNIQIGTLTETFDVDMHRAAGYTSNM 360
 Db 301 GDKVQIVGDDFFVTNPERLKEGIAKKAANSILVKVNIQIGTLTETFDVDMHRAAGYTSNM 360
 QY 361 SHRSGETEDTTIADLAVALNCGQIKTGAPASDRVAKYNOLLRIEQLLDAGVYAGRSAP 420
 Db 361 SHRSGETEDTTIADLAVALNCGQIKTGAPASDRVAKYNOLLRIEQLLDAGVYAGRSAP 420
 QY 421 PRFQ 425
 Db 421 PRFQ 425

RESULT 5

US-10-728-947-4
 ; Sequence 4, Application US/10728947
 ; Publication No. US20040220394A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Mockel, Bettina
 ; APPLICANT: Pfefferle, Walter
 ; APPLICANT: Hermann, Thomas
 ; APPLICANT: Pöhler, Alfred
 ; APPLICANT: Kalinowski, Jorn
 ; APPLICANT: Bathe, Brigitte
 ; TITLE OF INVENTION: New Nucleotide Sequences that Code for the Eno Gene
 ; FILE REFERENCE: 21123/278404
 ; CURRENT APPLICATION NUMBER: US/10/728,947
 ; CURRENT FILING DATE: 2003-12-08
 ; PRIOR APPLICATION NUMBER: US/09/860,768
 ; PRIOR FILING DATE: 2001-05-21
 ; NUMBER OF SEQ ID NOS: 6
 ; SOFTWARE: Patent in version 3.0
 ; SEQ ID NO 4
 ; LENGTH: 425
 ; TYPE: PRT
 ; ORGANISM: Corynebacterium glutamicum
 US-10-728-947-4

Query Match 99.8%; Score 2151; DB 16; Length 425;
 Best Local Similarity 99.8%; Pred. No. 5.6e-171;
 Matches 424; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VAEIMHVFAREILDSRGNTVFAEVLDDGSHGVAGVPSGASTGVHEAHELDDGDRYL 60
 Db 1 VAEIMHVFAREILDSRGNTVFAEVLDDGSHGVAGVPSGASTGVHEAHELDDGDRYL 60
 QY 61 KGVLKAVENNEIEIGDELAGLEADDDORLIDEAMIKLDGTANKSRILGANAILGVSMVAKA 120
 Db 61 KGVLKAVENNEIEIGDELAGLEADDDORLIDEAMIKLDGTANKSRILGANAILGVSMVAKA 120
 QY 121 AADSAGLPLFRYIGGPNHVLVPPMNIINGGAHDSGVVDVQEFMIAPITCAETTFSEALRN 180
 Db 121 AADSAGLPLFRYIGGPNHVLVPPMNIINGGAHDSGVVDVQEFMIAPITCAETTFSEALRN 180
 QY 181 GAEVTHALKSVIKKGLSTGLDGGFAPSVGSTREALDLIVEATEKAGTTPGKDIALAL 240
 Db 181 GAEVTHALKSVIKKGLSTGLDGGFAPSVGSTREALDLIVEATEKAGTTPGKDIALAL 240
 QY 241 DVASSEFFKDGTYHFEQGHSAEAMANYAELVDAYPIVSIEDPLQEDDWEQYTNLTATI 300
 Db 241 DVASSEFFKDGTYHFEQGHSAEAMANYAELVDAYPIVSIEDPLQEDDWEQYTNLTATI 300
 QY 301 GDKVQIVGDDFFVTNPERLKEGIAKKAANSILVKVNIQIGTLTETFDVDMHRAAGYTSNM 360
 Db 301 GDKVQIVGDDFFVTNPERLKEGIAKKAANSILVKVNIQIGTLTETFDVDMHRAAGYTSNM 360
 QY 361 SHRSGETEDTTIADLAVALNCGQIKTGAPASDRVAKYNOLLRIEQLLDAGVYAGRSAP 420
 Db 361 SHRSGETEDTTIADLAVALNCGQIKTGAPASDRVAKYNOLLRIEQLLDAGVYAGRSAP 420
 QY 421 PRFQ 425
 Db 421 PRFQ 425

RESULT 6
 US-10-781-014-72
 ; Sequence 72, Application US/10781014
 ; Publication No. US20040180408A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Pompejus, Markus
 ; APPLICANT: Kröger, Burkhard
 ; APPLICANT: Schroder, Hartwig
 ; APPLICANT: Zelder, Oskar
 ; APPLICANT: Haberhauser, Gregor
 ; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS INVOLVED IN CARBON METABOLISM AND ENERGY

; TITLE OF INVENTION: PRODUCTION
 ; FILE REFERENCE: BGI-126CPCN
 ; CURRENT APPLICATION NUMBER: US/10/781.014
 ; CURRENT FILING DATE: 2004-02-17
 ; PRIOR APPLICATION NUMBER: US 09/602,740
 ; PRIOR FILING DATE: 2000-06-23
 ; PRIOR APPLICATION NUMBER: 60/141,031
 ; PRIOR FILING DATE: 1999-06-25
 ; PRIOR APPLICATION NUMBER: 60/143,208
 ; PRIOR FILING DATE: 1999-07-09
 ; PRIOR APPLICATION NUMBER: 60/151,572
 ; PRIOR FILING DATE: 1999-08-31
 ; PRIOR APPLICATION NUMBER: DE 19931412.8
 ; PRIOR FILING DATE: 1999-07-08
 ; PRIOR APPLICATION NUMBER: DE 19931413.6
 ; PRIOR FILING DATE: 1999-07-08
 ; PRIOR APPLICATION NUMBER: DE 19931419.5
 ; PRIOR FILING DATE: 1999-07-08
 ; PRIOR APPLICATION NUMBER: DE 19931420.9
 ; PRIOR FILING DATE: 1999-07-08
 ; PRIOR APPLICATION NUMBER: DE 19931424.1
 ; PRIOR FILING DATE: 1999-07-08
 ; PRIOR APPLICATION NUMBER: DE 19931428.4
 ; PRIOR FILING DATE: 1999-07-08
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 784
 ; SEQ ID NO 72
 ; LENGTH: 425
 ; TYPE: PRT
 ; ORGANISM: Corynebacterium glutamicum
 ; US-10-781-014-72

Query Match 99.7%; Score 2149; DB 16; Length 425;
 Best Local Similarity 99.8%; Pred. No. 8.3e-171;
 Matches 424; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 VAEIMHVFAREILDSRGNPTVEAEVFLDDGSHGVAGVPSGASTGVHEAHLRDGGDRYLG 60
 DB 1 VAEIMHVFAREILDSRGNPTVEAEVFLDDGSHGVAGVPSGASTGVHEAHLRDGGDRYLG 60
 QY 61 KGVLKAVENVNEEIGDELGLADDDORLIDEAMIKLDGTANKSRKLGANAILGVSMVAKA 120
 DB 61 KGVLKAVENVNEEIGDELGLADDDORLIDEAMIKLDGTANKSRKLGANAILGVSMVAKA 120
 QY 121 AADSAGLPLFRYIGGNPAHVLVPVMNIIIGGAHADSGVDVQEFMIPIGAETFEALRN 180
 DB 121 AADSAGLPLFRYIGGNPAHVLVPVMNIIIGGAHADSGVDVQEFMIPIGAETFEALRN 180
 QY 181 GAENVHAKSVIKKGLSTGLDGGFAPS SVGSTREALDLIVEAIEKAGFTPGKDIALAL 240
 DB 181 GAENVHAKSVIKKGLSTGLDGGFAPS SVGSTREALDLIVEAIEKAGFTPGKDIALAL 240
 QY 241 DVASSEFFKDGTYHFEGGQHSAAEMANYAELVDAYPIVSIEDPLQEDDWEGYTNLTATI 300
 DB 241 DVASSEFFKDGTYHFEGGQHSAAEMANYAELVDAYPIVSIEDPLQEDDWEGYTNLTATI 300
 QY 301 GDKVQIVGDDPFVTNPERLKEGIAKKAANSILVKVNOIGTLTETFDVDMHARAGYTSMM 360
 DB 301 GDKVQIVGDDPFVTNPERLKEGIAKKAANSILVKVNOIGTLTETFDVDMHARAGYTSMM 360
 QY 361 SHRSGETEDTTIADLAVALNCGQIKTGAPASDRVAKYNQLLRIEQLLDGAGVYAGRSF 420
 DB 361 SHRSGETEDTTIADLAVALNCGQIKTGAPASDRVAKYNQLLRIEQLLDGAGVYAGRSF 420
 QY 421 PRFQG 425
 DB 421 PRFQG 425

RESULT 7
 US-10-494-836-8
 ; Sequence 8, Application US/10494836
 ; Publication No. US20050014233A1

; GENERAL INFORMATION:
 ; APPLICANT: Zelder, Oskar
 ; APPLICANT: Pompejus, Markus
 ; APPLICANT: Schroder, Hartwig
 ; APPLICANT: Kroger, Burkhard
 ; APPLICANT: Klopprogge, Corinna
 ; APPLICANT: Haberhauer, Gregor
 ; TITLE OF INVENTION: Genes coding for proteins of carbon metabolism and energy produc
 ; FILE REFERENCE: BGI-167US
 ; CURRENT APPLICATION NUMBER: US/10/494,836
 ; CURRENT FILING DATE: 2004-05-05
 ; PRIOR APPLICATION NUMBER: PCT/EP02/12135
 ; PRIOR FILING DATE: 2002-10-31
 ; PRIOR APPLICATION NUMBER: DE 101 54 270.4
 ; PRIOR FILING DATE: 2001-11-05
 ; NUMBER OF SEQ ID NOS: 116
 ; SEQ ID NO 8
 ; LENGTH: 425
 ; TYPE: PRT
 ; ORGANISM: Corynebacterium glutamicum
 ; US-10-494-836-8

Query Match 99.7%; Score 2149; DB 17; Length 425;

Best Local Similarity 99.8%; Pred. No. 8.3e-171;
 Matches 424; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 VAEIMHVFAREILDSRGNPTVEAEVFLDDGSHGVAGVPSGASTGVHEAHLRDGGDRYLG 60
 DB 1 VAEIMHVFAREILDSRGNPTVEAEVFLDDGSHGVAGVPSGASTGVHEAHLRDGGDRYLG 60
 QY 61 KGVLKAVENVNEEIGDELGLADDDORLIDEAMIKLDGTANKSRKLGANAILGVSMVAKA 120
 DB 61 KGVLKAVENVNEEIGDELGLADDDORLIDEAMIKLDGTANKSRKLGANAILGVSMVAKA 120
 QY 121 AADSAGLPLFRYIGGNPAHVLVPVMNIIIGGAHADSGVDVQEFMIPIGAETFEALRN 180
 DB 121 AADSAGLPLFRYIGGNPAHVLVPVMNIIIGGAHADSGVDVQEFMIPIGAETFEALRN 180
 QY 181 GAENVHAKSVIKKGLSTGLDGGFAPS SVGSTREALDLIVEAIEKAGFTPGKDIALAL 240
 DB 181 GAENVHAKSVIKKGLSTGLDGGFAPS SVGSTREALDLIVEAIEKAGFTPGKDIALAL 240
 QY 241 DVASSEFFKDGTYHFEGGQHSAAEMANYAELVDAYPIVSIEDPLQEDDWEGYTNLTATI 300
 DB 241 DVASSEFFKDGTYHFEGGQHSAAEMANYAELVDAYPIVSIEDPLQEDDWEGYTNLTATI 300
 QY 301 GDKVQIVGDDPFVTNPERLKEGIAKKAANSILVKVNOIGTLTETFDVDMHARAGYTSMM 360
 DB 301 GDKVQIVGDDPFVTNPERLKEGIAKKAANSILVKVNOIGTLTETFDVDMHARAGYTSMM 360
 QY 361 SHRSGETEDTTIADLAVALNCGQIKTGAPASDRVAKYNQLLRIEQLLDGAGVYAGRSF 420
 DB 361 SHRSGETEDTTIADLAVALNCGQIKTGAPASDRVAKYNQLLRIEQLLDGAGVYAGRSF 420
 QY 421 PRFQG 425
 DB 421 PRFQG 425

RESULT 8
 US-10-282-122A-53908
 ; Sequence 53908, Application US/10282122A
 ; Publication No. US20040029129A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Liangsu
 ; APPLICANT: Zamudio, Carlos
 ; APPLICANT: Malone, Cheryl
 ; APPLICANT: Haselbeck, Robert
 ; APPLICANT: Ohlsen, Kari
 ; APPLICANT: Zyskind, Judith
 ; APPLICANT: Wall, Daniel
 ; APPLICANT: Trawick, John
 ; APPLICANT: Carr, Grant


```

; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patcin version 3.1
; SEQ ID NO 62133
; LENGTH: 429
; TYPE: PRT
; ORGANISM: Mycobacterium avium
US-10-282-122A-62133

```

```

Query Match      72.3%; Score 1558.5; DB 15; Length 429;
Best Local Similarity 72.9%; Pred. No. 1.9e-121;
Matches 307; Conservative 46; Mismatches 67; Indels 1; Gaps 1;

Qy 4 IMHVFAREILDSRGNTVEAEVFLDDGSHGVAGVPGSGASTGVHEAHLRDGDRYLKGV 63
Db 4 IEQVGAREILDSRGNTVEAEVFLDDGSHGVAGVPGSGASTGVHEAHLRDGDRYLKGV 63

Qy 64 LKAVENVAEEIGDELGLAGLEADQRLIDEAMIKLDGTANKSRKLGANAILGVSMVAKAAD 123
Db 64 QKAVQAVLDEIGPAVIGLNADQRLVDQALVDLDTGPKSRKLGANAILGVSLAVAKAAD 123

Qy 124 SAGLPLFRYIGGNPAHVLVPMNNTINGGAHADSGVDVQEFMIAPIGAEFTSEALRNQAE 183
Db 124 SAEPLFRYIGGNPAHVLVPMNNTINGGAHADTAVIDIQEFMVAIPGAPSAEALRWGAE 183

Qy 184 VYHALKSIVKEKGLSTGLDGGEGFAPSVGSTREALDLIVEAEKAGFTPGKQIALADVA 243
Db 184 VYHSLSKIVKEKGLSTGLDGGEGFADVPAGTAAALDLIGRAIESAGFKLGTVDVALDAA 243

Qy 244 SSEFFKDTG - YHFEQGQHSAAEMANVYAEVLVDAYPIVSIEDPLQEDDWEYTNLTATIGD 302
Db 244 ATEFYSDGTGKFEQSTABQMAEFYAGLLGAYPLVSIEDPLSDDWDGWAALTASIGD 303

Qy 303 KQVIVGDDFFVTNPERLKEGIAKKAANSILVKVNOIGTLTETTFDVAVMHARAGVTSMMSH 362
Db 304 RVQLVGGDDVFTNPERLKEGIEKGVANALLVKVNOIGTLTETTLDAVALAHSHGVRVNMH 363

Qy 363 RSGETEDTTIADLAVALNCQIKTGAPARSDRVAKYNQLLRIEQLIGDAGVYAGRSAPPR 422
Db 364 RSGETEDTTIADLAVAGSGQIKTGAPARSKRVAKYNQLLRIEALGDAARYAGDLAPPR 423

Qy 423 F 423
Db 424 F 424

```

```

RESULT 11
US-10-156-761-11070
; Sequence 11070, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: HORIKAWA, JUN
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 11070
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-11070

Query Match      72.2%; Score 1555.5; DB 14; Length 428;
Best Local Similarity 73.2%; Pred. No. 3.4e-121;
Matches 311; Conservative 42; Mismatches 71; Indels 1; Gaps 1;

Qy 1 VAETMHVFAREILDSRGNTVEAEVFLDDGSHGVAGVPGSGASTGVHEAHLRDGQ-DRYL 59
Db 1 VPSIDVVVAREILDSRGNTVEAEVFLDDGSHGVAGVPGSGASTGVHEAHLRDGDPNRYQ 60

Qy 60 GKGVLKAVENVAEEIGDELGLAGLEADQRLIDEAMIKLDGTANKSRKLGANAILGVSMVAK 119
Db 61 GKGVEKAVLAVIEQIGPELVGYDATEQRLIDQAMFDLDATDNKGLGANAILGVSLAVAH 120

Qy 120 AADSDAGLPLFRYIGGNPAHVLVPMNNTINGGAHADSGVDVQEFMIAPIGAEFTSEALR 179
Db 121 AASEASDPLFRYIGGNPAHVLVPMNNTINGGAHADSNVDIQEFMIAPIGAEFTSEALR 180

Qy 180 NGAEVYHALKSIVKEKGLSTGLDGGEGFAPSVGSTREALDLIVEAEKAGFTPGKQIALA 239
Db 181 WGAEVYHTLAKVLTKGLSTGLDGGEGFAPNLSNRAALDLIIIEAKQAGYIFGEQIALA 240

Qy 240 LDVASEFFKDTGTHFEQGQHSAAEMANVYAEVLVDAYPIVSIEDPLQEDDWEYTNLTAT 299
Db 241 LDVAASEFFKDTGTHFEQGQHSAAEMTEYYEELVSAVPLVSIEDPLVDWAGWKVITDK 300

Qy 300 IGDKVQIVGDDFFVTNPERLKEGIAKKAANSILVKVNOIGTLTETTFDVAVMHARAGVTS 359
Db 301 LGDKVQIVGDDLFPVTNPERLARGIEGSAVALLVKNQIGSLTETTLDAVELAQRNGFKCM 360

Qy 360 MSHRSGETEDTTIADLAVALNCQIKTGAPARSDRVAKYNQLLRIEQLIGDAGVYAGRSA 419
Db 361 MSHRSGETEDTTIADLAVAVNCQIKTGAPARSDRVAKYNQLLRIEILLDDAAEYAGRSA 420

Qy 420 FPRFQ 424
Db 421 FPRFR 425

```

```

RESULT 12
US-10-282-122A-64490
; Sequence 64490, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wadg, Liangsu
; APPLICANT: Zamudio, Carlos

```

```

; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 64490
; LENGTH: 429
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-10-282-122A-64490

Query Match 72.0%; Score 1551.5; DB 15; Length 429;
Best Local Similarity 73.2%; Pred. No. 7.3e-121;
Matches 308; Conservative 41; Mismatches 71; Indels 1; Gaps 1;

Qy 4 IMHVFAREILDSRGNTVEAEVFLDDGSHGVAGVPSGASTGVHEAHELRDGGDRYLGKV 63
Db 4 IEQVFAREILDSRGNTVEAEVFLDDGSHGVAGVPSGASTGVHEAHELRDGGDRYLGKV 63
Qy 64 LKAVENNEEIGDELADGADDDORLIDEMIKLDTANKSRGLGANAILGVSMVAKAAAD 123
Db 64 QKAVQAVLDIGPAVIGLNADQRLVDQALVDLDTGTPDKSRGLGNAILGVSLAVAKAAAD 123
Qy 124 SAGLPLFRYIGGPNNAHILPVPMMNIIINGGAHDSGVQVDFMTAPIGAETPSEALRNGAE 183
Db 124 SAGLPLFRYIGGPNNAHILPVPMMNIIINGGAHDSGVQVDFMTAPIGAETPSEALRNGAE 183
Qy 363 RSGETEDTMIADLAVAGISGQIKTGAPARSERVAKYNQLLIEBALSDAARYAGDLAPPR 422
Db 363 RSGETEDTMIADLAVAGISGQIKTGAPARSERVAKYNQLLIEBALSDAARYAGDLAPPR 422

; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 64490
; LENGTH: 429
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-10-282-122A-64490

Query Match 72.0%; Score 1551.5; DB 15; Length 429;
Best Local Similarity 73.2%; Pred. No. 7.3e-121;
Matches 308; Conservative 41; Mismatches 71; Indels 1; Gaps 1;

Qy 4 IMHVFAREILDSRGNTVEAEVFLDDGSHGVAGVPSGASTGVHEAHELRDGGDRYLGKV 63
Db 4 IEQVFAREILDSRGNTVEAEVFLDDGSHGVAGVPSGASTGVHEAHELRDGGDRYLGKV 63
Qy 64 LKAVENNEEIGDELADGADDDORLIDEMIKLDTANKSRGLGANAILGVSMVAKAAAD 123
Db 64 QKAVQAVLDIGPAVIGLNADQRLVDQALVDLDTGTPDKSRGLGNAILGVSLAVAKAAAD 123
Qy 124 SAGLPLFRYIGGPNNAHILPVPMMNIIINGGAHDSGVQVDFMTAPIGAETPSEALRNGAE 183
Db 124 SAGLPLFRYIGGPNNAHILPVPMMNIIINGGAHDSGVQVDFMTAPIGAETPSEALRNGAE 183
Qy 363 RSGETEDTMIADLAVAGISGQIKTGAPARSDRVAKYNQLLIEOLLDGAGVYAGRSAPPR 422
Db 363 RSGETEDTMIADLAVAGISGQIKTGAPARSDRVAKYNQLLIEOLLDGAGVYAGRSAPPR 422

; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 62696
; LENGTH: 439
; TYPE: PRT
; ORGANISM: Mycobacterium bovis
US-10-282-122A-62696

Query Match 72.0%; Score 1551.5; DB 15; Length 439;
Best Local Similarity 73.2%; Pred. No. 7.6e-121;
Matches 308; Conservative 41; Mismatches 71; Indels 1; Gaps 1;

Qy 4 IMHVFAREILDSRGNTVEAEVFLDDGSHGVAGVPSGASTGVHEAHELRDGGDRYLGKV 63
Db 14 IEQVFAREILDSRGNTVEAEVFLDDGSHGVAGVPSGASTGVHEAHELRDGGDRYLGKV 73
Qy 64 LKAVENNEEIGDELADGADDDORLIDEMIKLDTANKSRGLGANAILGVSMVAKAAAD 123
Db 74 QKAVQAVLDIGPAVIGLNADQRLVDQALVDLDTGTPDKSRGLGNAILGVSLAVAKAAAD 133
Qy 124 SAGLPLFRYIGGPNNAHILPVPMMNIIINGGAHDSGVQVDFMTAPIGAETPSEALRNGAE 183
Db 134 SAGLPLFRYIGGPNNAHILPVPMMNIIINGGAHDSGVQVDFMTAPIGAETPSEALRNGAE 193
```

Qy 184 VYHAKSVIKKGLSTGLDGGGAPSVGSTRALDLIVEAEKAGFTPGKDIALDVA 243
 Db 194 VYHAKSVLKKEGLSTGLDGGGAPDVAGTAAALDLISRAIESAGLRPGADVALDAA 253
 Qy 244 SSEFFKQGT-YHFEQGQHSAAEMANVYAEVLVDVPIVSIEDPLOBDDWEGYTNLTATIGD 302
 Db 254 ATEFTDGTGVFEGTTFADQMTFYGALLGAYPLVSIEDPLOBDDWEGYTNLTATIGD 313
 Qy 303 KQVIGDDFFVTNPERLKEGIAKKAANSILVKVNOIGTLTETFDVADMAHRAGYTSMMSH 362
 Db 314 RVQVIGDDIFVTNPERLEEGIERGVANALLVKVNOIGTLTETDVAHLAHHGGYRTWISH 373
 Qy 363 RSGETEDTTIADLAVALNCGQIKTGAPARSDRVAKYNQILLRIEQLLGDAGVYAGRSAPPR 422
 Db 374 RSGETEDTTIADLAVAIAGSQIKTGAPARSDRVAKYNQILLRIEALGDAARYAGDLAPPR 433
 Qy 423 F 423
 Db 434 F 434

RESULT 14
 US-10-282-122A-63682
 ; Sequence 63682, Application US/10282122A
 ; Publication No. US20040029129A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Liangsu
 ; APPLICANT: Zamudio, Carlos
 ; APPLICANT: Malone, Cheryl
 ; APPLICANT: Haselbeck, Robert
 ; APPLICANT: Ohlsen, Kari
 ; APPLICANT: Zyskind, Judith
 ; APPLICANT: Wall, Daniel
 ; APPLICANT: Trawick, John
 ; APPLICANT: Carr, Grant
 ; APPLICANT: Yamamoto, Robert
 ; APPLICANT: Forsyth, R.
 ; APPLICANT: Xu, H.
 ; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
 ; FILE REFERENCE: ELITRA.034A
 ; CURRENT APPLICATION NUMBER: US/10/282,122A
 ; CURRENT FILING DATE: 2003-02-20
 ; PRIOR APPLICATION NUMBER: 60/191,078
 ; PRIOR FILING DATE: 2000-03-21
 ; PRIOR APPLICATION NUMBER: 60/206,848
 ; PRIOR FILING DATE: 2000-05-23
 ; PRIOR APPLICATION NUMBER: 60/207,727
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: 60/230,335
 ; PRIOR FILING DATE: 2000-09-06
 ; PRIOR APPLICATION NUMBER: 60/230,347
 ; PRIOR FILING DATE: 2000-09-09
 ; PRIOR APPLICATION NUMBER: 60/242,578
 ; PRIOR FILING DATE: 2000-10-23
 ; PRIOR APPLICATION NUMBER: 60/253,625
 ; PRIOR FILING DATE: 2000-11-27
 ; PRIOR APPLICATION NUMBER: 60/257,931
 ; PRIOR FILING DATE: 2000-12-22
 ; PRIOR APPLICATION NUMBER: 60/267,636
 ; PRIOR FILING DATE: 2001-02-09
 ; PRIOR APPLICATION NUMBER: 60/269,308
 ; PRIOR FILING DATE: 2001-02-16
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 78614
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 63682
 ; LENGTH: 447
 ; TYPE: PRT
 ; ORGANISM: Mycobacterium leprae

Query Match 69.7%; Score 1501.5; DB 15; Length 447;
 Best Local Similarity 70.8%; Pred. No. 1.2e-116;

Matches 300; Conservative 47; Mismatches 76; Indels 1; Gaps 1;
 Qy 1 VAEIMHVAREILDSRGNPTVEAEVFLDDGSHGVAGVPSGASTGVHBAHELRDGDRYLG 60
 Db 19 VPVIEQVAREILDSRGNPTVEVEVLIDGTFFARAAVPSGASTGEYEAVELRDGDRYGG 78
 Qy 61 KGVLKAVENTNEBIGDELADDDQRLIDEAMIKLDTANKSRGLGANAILGVSMAYAKA 120
 Db 79 KGVKRAVDAVLDRIGPVVIGLNANDQRLIDQELLDLDTGTPKSRGLGANAILGVSLAVAKA 138
 Qy 121 AADSAGLPLFRYTGPNNAHVLPVPMNMIINGGAHADSGVDQBPMAPIGAEFTSEALRN 180
 Db 139 AADSAGLPLFRYTGPNNAHVLPVPMNMIINGGAHADTAADVQBPMAPIGAEFTSEALRN 198
 Qy 181 GAETHYALKSVIKKGLSTGLDGGGAPSVGSTRALDLIVEAEKAGFTPGKDIALAL 240
 Db 199 GAETHYALKSVLKKGKGLSTGLDGGGAPFVAGTTAALDLVTLAIEAAGFKPGADVALAL 258
 Qy 241 DVASSEFFKQGT-YHFEQGQHSAAEMANVYAEVLVDVPIVSIEDPLOBDDWEGYTNLTAT 299
 Db 259 DAAATEFTDGTGVFEGTTFADQMTFYGALLGAYPLVSIEDPLOBDDWEGYTNLTAT 318
 Qy 300 IGDKQVIGDDFFVTNPERLKEGIAKKAANSILVKVNOIGTLTETFDVADMAHRAGYTSM 359
 Db 319 IGEQVQVIGDDIFATNPERLEEGIRGVANALLVKVNOIGTLTETLEVALAHHSGYRTM 378
 Qy 360 MSHRSGETEDTTIADLAVALNCGQIKTGAPARSDRVAKYNQILLRIEQLLGDAGVYAGRSA 419
 Db 379 ISHRSGETEDTTIADLAVALGSGQIKTGAPARSDRVAKYNQILLRIEELGDAARYAGDLA 438
 Qy 420 FPRF 423
 Db 439 FLRY 442

RESULT 15
 US-10-369-493-11696
 ; Sequence 11696, Application US/10369493
 ; Publication No. US20030233675A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Chen, Xianfeng
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
 ; FILE REFERENCE: 38-10(54052)B
 ; CURRENT APPLICATION NUMBER: US/10/369,493
 ; CURRENT FILING DATE: 2003-02-28
 ; PRIOR APPLICATION NUMBER: US 60/360,039
 ; PRIOR FILING DATE: 2002-02-21
 ; NUMBER OF SEQ ID NOS: 47374
 ; SEQ ID NO 11696
 ; LENGTH: 423
 ; TYPE: PRT
 ; ORGANISM: Agrobacterium tumefaciens
 US-10-369-493-11696

Query Match 65.8%; Score 1418; DB 15; Length 423;
 Best Local Similarity 65.9%; Pred. No. 1e-109;
 Matches 275; Conservative 59; Mismatches 81; Indels 2; Gaps 1;
 Qy 4 IMHVAREILDSRGNPTVEAEVFLDDGSHGVAGVPSGASTGVHBAHELRDGDRYLGKV 63
 Db 4 ITDIAAREILDSRGNPTVEVDVLEDDGSGRAAVPSGASTGAHEAELRDGDRYLGKV 63
 Qy 64 LKAVENYNEBIGDELADDDQRLIDEAMIKLDTANKSRGLGANAILGVSMAYAKAAD 123
 Db 64 EKAVEAVNTEIFDAIGGFDAAENQIQIDQMMIALDGTENKSRGLGANAILGVSLATAKAAAE 123
 Qy 124 SAGLPLFRYTGPNNAHVLPVPMNMIINGGAHADSGVDQBPMAPIGAEFTSEALRNAGAE 183
 Db 124 SAGLPLFRYTGPNNAHVLPVPMNMIINGGAHADSGVDQBPMAPIGAEFTSEALRNAGAE 183


```

Db 124 ASGLPLYRYVGGPNAHLFPVPMNIIINGGAHADNPIDQOFMILPVGAENIREAVRMGSE 183
Qy 184 VYHALKSVIKEGLSTGLDEGGFAPSVGSTREALDLIVEAIEKAGFTPGKDIALALDVA 243
Db 184 VFHTLKEISAQGHNTNVGDEGGFAPGLESAPALDFIMKSIKAGIRPEEDMYGLDCA 243
Qy 244 SSEFFKOGTYHPEGGQHS--AAEMANVYAELVDAYPIVSIEDPLQEDDWEGYTNLTATIG 301
Db 244 STEFFKOGKYVLEGEGRITLEPGAMAEBYLAELVNKYPISVEDGMAEDDWEGWKTLTDLVG 303
Qy 302 DKVOIVGDDFFVTNPERLKEGIKKAANSILVKVNOIGTITFTFDVDMHAHRAGYTSMS 361
Db 304 NKQOLVGDDLFVTNSARLRDGINMGVANSILVKVNOIGSLSETLDAVETAHKAGYTAVMS 363
Qy 362 HRSGETEDTTIADLAVALNCGOIKTGAPARSDRVAKYNOLLRIEOLLGDAGVYAGRS 418
Db 364 HRSGETEDTTIADLAVATNCGOIKTGSLARSRLAKYNQLIRIEMLGPOAYAGRS 420

```

Search completed: September 29, 2005, 14:45:39
Job time : 1028 secs

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 28, 2005, 16:02:19 ; Search time 6707 Seconds
(without alignments)
11400.377 Million cell updates/sec

Title: US-10-728-947-3

Perfect score: 1578

Sequence: 1 ggcctggggtatggtagttt.....ctcaagcagggagcgtgctt 1578

Scoring table: IDENTITY_NUC.

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_htg.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1578	100.0	1578	6	AR490886
2	1576.4	99.9	1578	6	AR490885 Sequence
3	1576.4	99.9	1578	6	AX136862 Sequence
4	1576.4	99.9	1578	6	BD014844 Polynucle
5	1576.4	99.9	333150	1	AP005277 Coryneb
6	1576.4	99.9	348475	1	AX927150 Coryneb
7	1576.4	99.9	349980	6	AX127146 Sequence
8	1401.8	88.8	1405	6	AX763190 Sequence
9	1394.8	88.4	1398	6	AX064945 Sequence
10	1273.4	80.7	1275	6	BD163286 Novel pol
11	1273.4	80.7	1275	6	AX121169 Sequence
12	1034.8	65.6	300750	1	AP005217 Coryneb
13	915.8	58.0	347625	1	BX248356 Coryneb
14	732.4	46.4	276800	1	SC0939115 Streptomy
15	700	44.4	110000	1	AX016822_17 Continuation (18 o
16	696.8	44.2	296300	1	AP005035 Streptomy
17	680.6	43.1	110000	1	AP006618_50 Continuation (51 o
18	657.6	41.7	303855	1	AE017230 Mycobacte
19	640.8	40.6	30159	6	CQ363776 Sequence

20	640.8	40.6	110000	1	AE017283_05 Continuation (6 of
21	632.6	40.1	45191	6	AX704278 Sequence
22	632.6	40.1	110000	1	AE000516_11 Continuation (12 o
23	632.6	40.1	327650	1	BX248337 Mycobacte
24	632.6	40.1	349306	1	BX842575 Mycobacte
25	590	37.4	329100	1	SMES591787
26	583.4	37.0	344050	1	MLEPRTN1
27	583	36.9	10029	1	AE014725 Bifidobac
28	583	36.9	349980	6	AX492784 Sequence
29	583	36.9	349980	6	AX492785 Sequence
30	583	36.9	349980	6	AX553951 Sequence
31	583	36.9	349980	6	AX553952 Sequence
32	579.4	36.7	12058	1	AE009103 Agrobacte
33	579.4	36.7	12407	1	AE008068 Caulobact
34	578	36.6	10934	1	AE005846 Mesorhizo
35	575.4	36.5	347660	1	AP002394 Brucella
36	574.8	36.4	12709	1	AE014414 Brucella
37	571.6	36.2	13776	1	AE009526 Methyloba
38	570.2	36.1	1290	1	AX181039 Rhodopseu
39	569.8	36.1	349981	1	BX572602 Rhodopseu
40	559.6	35.5	300600	1	AP005952 Bradyrhiz
41	557.8	35.3	193050	1	AL646062 Ralstonia
42	555.6	35.2	197615	5	BX470161 Zebraphish
43	550.2	34.9	110000	1	AE017180_24 Continuation (25 o
44	550.2	34.9	110000	1	AE017180_25 Continuation (26 o
45	542.8	34.4	110000	1	BX571965_27 Continuation (28 o

ALIGNMENTS

RESULT 1	AR490886	Sequence 3 from patent US 6713289.	1578 bp	DNA	linear	PAT 15-MAY-2004
LOCUS	AR490886					
DEFINITION	AR490886					
ACCESSION	AR490886					
VERSION	AR490886.1	GI:47258405				
KEYWORDS	Unknown.					
SOURCE	Unknown.					
ORGANISM	Unknown.					
REFERENCE	1 (bases 1 to 1578)					
AUTHORS	Mockel,B., Pfefferle,W., Hermann,T., Puhler,A., Kalinowski,J. and Bathe,B.					
TITLE	Nucleotide sequences which code for the eno gene					
JOURNAL	Patent: US 6713289-A 3 30-MAR-2004;					
FEATURES	Location/Qualifiers					
source	1..1578					
	/organism="unknown"					
	/mol_type="genomic DNA"					

ORIGIN	Query Match	100.0%	Score 1578;	DB 6;	Length 1578;
	Best Local Similarity	100.0%;	Pred. No. 3.4e-289;		
	Matches 1578;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	GGCTGGGGATATGGTAGTTTTCGCCCACTAATTTCAACTGATTCCTCATCGAACACAGA	60		
Db	1	GGCTGGGGATATGGTAGTTTTCGCCCACTAATTTCAACTGATTCCTCATCGAACACAGA	60		
Qy	61	TTCTGTCACAAATTTGGGTGTAGCTGATTTGAAGACATTTTGCATCAGTGAATTAATTTCTAG	120		
Db	61	TTCTGTCACAAATTTGGGTGTAGCTGATTTGAAGACATTTTGCATCAGTGAATTAATTTCTAG	120		
Qy	121	TTAGTCTCCCAAGTTGGCATAGGAGCCACAGTGGCTGAAATCATGCAGTATTCGCTCGC	180		
Db	121	TTAGTCTCCCAAGTTGGCATAGGAGCCACAGTGGCTGAAATCATGCAGTATTCGCTCGC	180		
Qy	181	GAATTTCTCGACTCCCGGGTAAACCAACCGTCGAGCAGAGGTTTTCTTGGATGACGGT	240		
Db	181	GAATTTCTCGACTCCCGGGTAAACCAACCGTCGAGCAGAGGTTTTCTTGGATGACGGT	240		
Qy	241	TCCACCGGTGTGGCAGGTGTTTCCATCCGCGGCATCCACCGCGCTCCACGAGGCTCATGAG	300		

Db 421 GAAGCAATGATCAAGCTTGGATGACCGGCAACAGTCCGCGCTGGGTGCAACGCAATC 480
Qy 481 CTTGTGTTTCCATGGCTGTGCAAAAGGCTGCTGCTGATTCGCGAGGCTCCCACTGTTTC 540
Db 481 CTTGTGTTTCCATGGCTGTGCAAAAGGCTGCTGCTGATTCGCGAGGCTCCCACTGTTTC 540
Qy 541 CGTACATCGGTGAGCAAAACGACAGCTTCTCCAGTTCCAGTTCCAAATGATGAACATCAATCAAC 600
Db 541 CGTACATCGGTGAGCAAAACGACAGCTTCTCCAGTTCCAGTTCCAAATGATGAACATCAATCAAC 600
Qy 601 GTGGGGCTCAGCTGACCTCGGCTGTGAGCTTCAGGAAATCATGATCGCTCCAAATCGGT 660
Db 601 GTGGGGCTCAGCTGACCTCGGCTGTGAGCTTCAGGAAATCATGATCGCTCCAAATCGGT 660
Qy 661 GCAGAGACCTTCTCTGAGGCTCTCCGCAACGGCGGAGGTCTACCAACGACTGAAGTCC 720
Db 661 GCAGAGACCTTCTCTGAGGCTCTCCGCAACGGCGGAGGTCTACCAACGACTGAAGTCC 720
Qy 721 GTCATCAAGAAAGGCTGTGTCACCGGACTTGGCGATGAGGGGGCTTCGCTCTCTCC 780
Db 721 GTCATCAAGAAAGGCTGTGTCACCGGACTTGGCGATGAGGGGGCTTCGCTCTCTCC 780
Qy 781 GTCGCTCCACCGCTGAGGCTCTTGACCTTATGCTTGGCAATCGAGAGGCTGGCTTC 840
Db 781 GTCGCTCCACCGCTGAGGCTCTTGACCTTATGCTTGGCAATCGAGAGGCTGGCTTC 840
Qy 841 ACCCAGGCAAGACATCGCTCTGCTGCGAGCTTCTCTGAGTCTTCAAGGAC 900
Db 841 ACCCAGGCAAGACATCGCTCTGCTGCGAGCTTCTCTGAGTCTTCAAGGAC 900
Qy 901 GGCACCTACCACTTCCGAGGTGGCAGCACTCCGAGCTGAGATGCGCAACGTTTACGCT 960
Db 901 GGCACCTACCACTTCCGAGGTGGCAGCACTCCGAGCTGAGATGCGCAACGTTTACGCT 960
Qy 961 GAGCTGTTGACGCTGACCAATGCTCTCATGAGGACCCACTGCGAGGAGATGACTGG 1020
Db 961 GAGCTGTTGACGCTGACCAATGCTCTCATGAGGACCCACTGCGAGGAGATGACTGG 1020
Qy 1021 GAGGTTACCAACCTCACCAGCAACCTACCGCAACAGTTCAGATGCTTGGCGACAC 1080
Db 1021 GAGGTTACCAACCTCACCAGCAACCTACCGCAACAGTTCAGATGCTTGGCGACAC 1080
Qy 1081 TTCTTCGTCAACACCTGAGCGCTGAGGAGGCGATCGCTAAGAGGCTCCCACTCC 1140
Db 1081 TTCTTCGTCAACACCTGAGCGCTGAGGAGGCGATCGCTAAGAGGCTCCCACTCC 1140
Qy 1141 ATCTGTTTAAAGTGAAACAGATCGGTACCTTCAACGAGACCTTCGACGCTGCGACATG 1200
Db 1141 ATCTGTTTAAAGTGAAACAGATCGGTACCTTCAACGAGACCTTCGACGCTGCGACATG 1200
Qy 1201 GCTACCGGCGAGGCTACACCTCCATGATGTCCTCCACCGTTCGCGTGAGACCGAGACAC 1260
Db 1201 GCTACCGGCGAGGCTACACCTCCATGATGTCCTCCACCGTTCGCGTGAGACCGAGACAC 1260
Qy 1261 ACCATTGCTGACCTCGCAGTTGCACTCAACTGTGCGCAGATCAAGACTGCTCCAGCA 1320
Db 1261 ACCATTGCTGACCTCGCAGTTGCACTCAACTGTGCGCAGATCAAGACTGCTCCAGCA 1320
Qy 1321 CGTTCGCAACCGGTGCGCAAGTACCAACGAGTTCCTCCGATCAGCAGCTGCTTGGCGAC 1380
Db 1321 CGTTCGCAACCGGTGCGCAAGTACCAACGAGTTCCTCCGATCAGCAGCTGCTTGGCGAC 1380
Qy 1381 GCGGCGCTTACGAGGTGCGAGCGATTCACCGCTTTCAGGGCTTAATAAAGCGCTT 1440
Db 1381 GCGGCGCTTACGAGGTGCGAGCGATTCACCGCTTTCAGGGCTTAATAAAGCGCTT 1440
Qy 1441 TTCAGCGCGGTACCTCAAGTTGCGGGGCTGTTGCGCTTACTACTGTTACTGCTGT 1500
Db 1441 TTCAGCGCGGTACCTCAAGTTGCGGGGCTGTTGCGCTTACTACTGTTACTGCTGT 1500
Qy 1501 GACTATGATCGAGGATTTATGGCAAGCAGAGAAACTCATAAAGGCTTGTTCCTGCTCT 1560

1501 GACTATGATCGAGGATTTATGGCAAGCAGAGAAACTCATAAAGGCTTGTTCCTGCTCT 1560
Qy 1561 CAAGCAGGCAACGCTGCTT 1578
Db 1561 CAAGCAGGCAACGCTGCTT 1578
RESULT 3
AX136862
LOCUS AX136862 1578 bp DNA linear PAT 30-MAY-2001
DEFINITION Sequence 1 from Patent EP1090998.
ACCESSION AX136862
VERSION AX136862.1 GI:14273213
SOURCE Corynebacterium glutamicum
ORGANISM Corynebacterium glutamicum
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Corynebacteriaceae; Corynebacterium.
REFERENCE 1
AUTHORS Moeckel, B., Pfefferle, W., Hermann, T., Puehler, A., Kalinowski, J. and Bathe, B.
TITLE Nucleotide sequences coding for the eno gene
JOURNAL Patent: EP 1090998-A 1 11-APR-2001;
Degussa AG (DE)
FEATURES
Location/Qualifiers
source 1. 1578
/organism="Corynebacterium glutamicum"
/mol_type="unassigned DNA"
/db_xref="taxon:1718"
151..1428
/note="unnamed protein product"
/codon_start=1
/transl_table=1
/protein_id="CAC39860.1"
/db_xref="GI:14273214"
/translation="MAIMHVFAREIILDSRGNPTVEAEVFLDDSGHGVGVPSSGASTG
VHEAHLRDGDRYLGKVLKAVENNEEIGDELAGEADQRLIDEAMIKLDGTANK
SRLCANAILGVSMAYAKAASAGLPLFRYIGGNHVLVPMWNI INGAHAGSGVD
VORPMIPIGAETSEALRNGAEVYHAKSVI KEKSLSTGLGDSGFPAPSVGSTREAL
DLIVEALEKAGFTPKDIALALDVAASSEFFKDKTYHFEFGQHSAAEMANYAEVLVAY
PVIQEDTFLQEDDWEVYTNLTATIGDKVQIVGDDFFVTNPERLKEGIAKKAANSILVK
VNIQTLETDFDAVDMHARAGTSMHRSGETEDTTIADLAVALNCQIKTGAPARS
DRVAKYNQLLRLEIQLLDAGVYAGRSAPFRPQG"
ORIGIN
Query Match 99.9%; Score 1576.4; DB 6; Length 1578;
Best Local Similarity 99.9%; Pred. No. 6.8e-289;
Matches 1577; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 GGCTGGGGATATGGGTAGTGTTCGCCCACTAATTTCACTGATTCCTCATCGAACACAGA 60
Db 1 GGCTGGGGATATGGGTAGTGTTCGCCCACTAATTTCACTGATTCCTCATCGAACACAGA 60
Qy 61 TTCTGTCGCAACAAATTTGGGTAGACGTGATTGAAGACATTTTCATCACGTGAATAATTTCTAG 120
Db 61 TTCTGTCGCAACAAATTTGGGTAGACGTGATTGAAGACATTTTCATCACGTGAATAATTTCTAG 120
Qy 121 TTAGCTCCCAAGTTGGCATAGGAGGCCACAGTGGCTGAAATCATGACGTATTCGCTCGC 180
Db 121 TTAGCTCCCAAGTTGGCATAGGAGGCCACAGTGGCTGAAATCATGACGTATTCGCTCGC 180
Qy 181 GAAATTTCTGACCTCCCGGCTAACCCAAACCGTCGAGGAGAGGTTTTCCTGGATGACGGT 240
Db 181 GAAATTTCTGACCTCCCGGCTAACCCAAACCGTCGAGGAGAGGTTTTCCTGGATGACGGT 240
Qy 241 TCCCAACCGGTGTGCGAGGTGTTCATCCGGGCGCATCCACCGGCTCCACGAGGCTCATGAG 300
Db 241 TCCCAACCGGTGTGCGAGGTGTTCATCCGGGCGCATCCACCGGCTCCACGAGGCTCATGAG 300
Qy 301 CTGCGTGACGGTGGCGATCGCTACCTGGGCAAGGGCGCTTTTGAAGGAGGTTGAAAGAGTC 360
Db 301 CTGCGTGACGGTGGCGATCGCTACCTGGGCAAGGGCGCTTTTGAAGGAGGTTGAAAGAGTC 360

Qy	361	AACGAAAGAAATCGCGCAGCAGCTCGCTGGCCCTAGAGGCTGACGATCAGCGCCTCATCGAC	420
Db	361	AACGAAAGAAATCGCGCAGCAGCTCGCTGGCCCTAGAGGCTGACGATCAGCGCCTCATCGAC	420
Qy	421	GAAAGCAATGATCAAGCTTTGATGGCACCGCCAAACAGTCCCGCTGGGGTGCAAACGGCAATC	480
Db	421	GAAAGCAATGATCAAGCTTTGATGGCACCGCCAAACAGTCCCGCTGGGGTGCAAACGGCAATC	480
Qy	481	CTTCGGTGTTCATATGGCTGTGTCAAAGGCTGCTGCTGATTCGCGACAGCCCTCCCACTGTTC	540
Db	481	CTTCGGTGTTCATATGGCTGTGTCAAAGGCTGCTGCTGATTCGCGACAGCCCTCCCACTGTTC	540
Qy	541	CGCTACATCGGTGACCAACCGCAACGACACGCTTCTTCAGGTTCCAAATGATGAACATCATCAAC	600
Db	541	CGCTACATCGGTGACCAACCGCAACGACACGCTTCTTCAGGTTCCAAATGATGAACATCATCAAC	600
Qy	601	GGTGGCGCTACGCTGACCTCCGGTGTGACGTTCAGGAAATTCATGATCGCTCCCAATCGGT	660
Db	601	GGTGGCGCTACGCTGACCTCCGGTGTGACGTTTTCAGGAAATTCATGATCGCTCCCAATCGGT	660
Qy	661	GCAGAGACCTTCTCTGAGGCTCTCCGCAACGGCGCGAGGTCTACCAACGACATGAAGTCC	720
Db	661	GCAGAGACCTTCTCTGAGGCTCTCTCGCAACGGCGCGAGGTCTACCAACGACATGAAGTCC	720
Qy	721	GTCAATCAAGGAAAGGGCCTGTCCACCGGACTTGGCGATGAGGCGGCTTGCTCTCTTC	780
Db	721	GTCAATCAAGGAAAGGGCCTGTCCACCGGACTTGGCGATGAGGCGGCTTGCTCTCTCTTC	780
Qy	781	GTCCGCTCCACCGTGAGGCTCTTGAACCTTATCGTTAAGGCAATCGAAGGCTGCGCTTC	840
Db	781	GTCCGCTCCACCGTGAGGCTCTTGAACCTTATCGTTGAGGCAATCGAAGGCTGCGCTTC	840
Qy	841	ACCCAGGCAAGGACATCGCTCTTGCTCTGACGTTGCTTCTCTGAGTTCTTCAAGGAC	900
Db	841	ACCCAGGCAAGGACATCGCTCTTGCTCTGACGTTGCTTCTCTGAGTTCTTCAAGGAC	900
Qy	901	GGCACCTACCACTTCGAAGGTGGCCAGCACCTCCGACGTGAGATGGCAAACTGTTACGCT	960
Db	901	GGCACCTACCACTTCGAAGGTGGCCAGCACCTCCGACGTGAGATGGCAAACTGTTACGCT	960
Qy	961	GAGCTCGTTGACGGGTACCCAAATCGTCTCATTGAGGACCCACATGCGAGGAAGATGATGG	1020
Db	961	GAGCTCGTTGACGGGTACCCAAATCGTCTCATTGAGGACCCACATGCGAGGAAGATGATGG	1020
Qy	1021	GAGGTTTACACCAACTCACCAGCAACATCGCGCAACAGGTTCAGATGCTGTTGGCGAGCAC	1080
Db	1021	GAGGTTTACACCAACTCACCAGCAACATCGCGCAACAGGTTCAGATGCTGTTGGCGAGCAC	1080
Qy	1081	TTCTTCGTTCACCAACCCCTGAGCGGCTGAAGGAGGGCATCGCTAAGAAGGCTGCCAACTCC	1140
Db	1081	TTCTTCGTTCACCAACCCCTGAGCGGCTGAAGGAGGGCATCGCTAAGAAGGCTGCCAACTCC	1140
Qy	1141	ATCTCGTTTAAAGTGAAACAGATCGGTACCTCACCAGACACTTTCGACGCTGTGCAATG	1200
Db	1141	ATCTCGTTTAAAGTGAAACAGATCGGTACCTCACCAGACACTTTCGACGCTGTGCAATG	1200
Qy	1201	GCTCACCAGGAGGCTACACTCCATGATGTCACCGTTCCGCTGAGACCGAGGACCC	1260
Db	1201	GCTCACCAGGAGGCTACACTCCATGATGTCACCGTTCCGCTGAGACCGAGGACCC	1260
Qy	1261	ACCAATGCTGACCTCGCAGTTGCACTCAACTGTGGCCAGATCAAGACTGGTGTCCAGCA	1320
Db	1261	ACCAATGCTGACCTCGCAGTTGCACTCAACTGTGGCCAGATCAAGACTGGTGTCCAGCA	1320
Qy	1321	CGTTCCGACCGGTGCGCAAAAGTACAAACAGCTTCTCCGCAATCGAGCAGCTGCTTGGCGAC	1380
Db	1321	CGTTCCGACCGGTGCGCAAAAGTACAAACAGCTTCTCCGCAATCGAGCAGCTGCTTGGCGAC	1380
Qy	1381	GCCGGCGTCTACGAGGTGCGAGCGATTTCCACGCTTTTCAGGGCTAAATTAAGACGCTT	1440
Db	1381	GCCGGCGTCTACGAGGTGCGAGCGATTTCCACGCTTTTCAGGGCTAAATTAAGACGCTT	1440
Qy	1441	TTGCAAGCCCGGTAACTCAAGGTTGCGGGCGCTGTTGCTTACTACTGTTACTGCTGT	1500

Db	1441	TTTCAGCGCCGGTAACCTCAAGCTTCGCGGGCGTGGTGGCTTTACTACTGTACTGTGT	1500
Qy	1501	GACTATGATCGAGATTATGGCAAGACAGAAAGAAACTCATAAAGCGCTTGTTCCTGTCT	1560
Db	1501	GACTATGATCGAGATTATGGCAAGACAGAAAGAAACTCATAAAGCGCTTGTTCCTGTCT	1560
Qy	1561	CAAGCAGGGGAACGTGCTT	1578
Db	1561	CAAGCAGGGGAACGTGCTT	1578
RESULT 4			
LOCUS	BD014844		
DEFINITION	Polynucleotide, DNA, isolation of polynucleotide sequence, and method of fermentative production of L-amino acid.		
ACCESSION	BD014844		
VERSION	BD014844.1	GI:22555651	
KEYWORDS	JP 2001161380-A/1.		
SOURCE	Corynebacterium glutamicum		
ORGANISM	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacterineae; Corynebacteriaceae; Corynebacterium.		
REFERENCE	1	(bases 1 to 1578)	
AUTHORS	Meckel, V., Pfeifferle, W., Hermann, T., Puehler, A., Kalinovski, J. and Baate, B.		
TITLE	Polynucleotide, DNA, isolation of polynucleotide sequence, and method of fermentative production of L-amino acid		
JOURNAL	Patent: JP 2001161380-A 1 19-JUN-2001;		
COMMENT	DEGUSSA HUELS AG		
	OS	Corynebacterium glutamicum	
	PN	JP 2001161380-A/1	
	PD	19-JUN-2001	
	PF	04-OCT-2000 JP 2000305110	
	PR	05-OCT-1999 DE 19947791.4	
	PI	VERTYNA MECKEL, WALTER PFEIFFERLE, THOMAS HERMANN, ALFRED PUEHLER,	
	PI	JOERN KALINOVSKI, BRIGITTE BAATE	
	PC	C12N15/09, C12N1/21, C12N9/88, C12P13/04, C12P13/08/(C12N15/09,	
	PC	C12R1:15).	
	PC	(C12P13/04, C12R1:15), (C12P13/08, C12R1:15), C12N15/00, (C12N15/00, PC	
		C12R1:15)	
	CC	Polynucleotide, DNA, isolation of polynucleotide sequence, and	
	CC	method of	
	CC	fermentative production of L-amino acid	
	FH	Key Location/Qualifiers	
	FT	CDS (151)..(1425).	
FEATURES			
source	1..1578	Location/Qualifiers	
	/organism="Corynebacterium glutamicum"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:1718"		
ORIGIN			
Query Match	99.9%	Score 1576.4	DB 6; Length 1578;
Best Local Similarity	99.9%	Pred. No. 6.8e-289;	
Matches 1577;	Conservative	0; Mismatches	1; Indels 0; Gaps 0;
Qy	1	GGCTGGGGATATGGGTAGTTTTCGCCACTAATTTCAACTGATTCCTCATCGAAACAAGA	60
Db	1	GGCTGGGGATATGGGTAGTTTTCGCCACTAATTTCAACTGATTCCTCATCGAAACAAGA	60
Qy	61	TTCTGTCAACAATTGGGTAGACGTGATTGAAGACATTTGATCACGTAATATTTCTAG	120
Db	61	TTCTGTCAACAATTGGGTAGACGTGATTGAAGACATTTGATCACGTAATATTTCTAG	120
Qy	121	TTAGTCTCCCAAGTTGGCATAGGAGGCCACAGTGGCTGAATCATGCACGATTTCGCTCGC	180
Db	121	TTAGTCTCCCAAGTTGGCATAGGAGGCCACAGTGGCTGAATCATGCACGATTTCGCTCGC	180
Qy	181	GAATTTCTCGACTCCCGCGGTAAACCCAAACCGTCGAGGACAGGTTTTTCTCGATGACGGT	240

181 GAAATTTCTCGACTCCGCGGTAAACCAACCGTGCAGGAGGTTTCTCGATGACGCT 240
241 TCCACCGGTGTCGAGGTGTTCCATCCGGCGCATCCACCGGCGTCCACGAGGCTCATGAG 300
241 TCCACCGGTGTCGAGGTGTTCCATCCGGCGCATCCACCGGCGTCCACGAGGCTCATGAG 300
301 CTGCGTGAACGGTGGCGATCGTACCTCGGCAAGGGCGTTTGAAGGCGAGTTGAAAACGTC 360
301 CTGCGTGAACGGTGGCGATCGTACCTCGGCAAGGGCGTTTGAAGGCGAGTTGAAAACGTC 360
361 AACGAGAATCGGCGACGAGCTCGTGGCTAGAGGCTGACGATCAGGCGCTCATCGAC 420
361 AACGAGAATCGGCGACGAGCTCGTGGCTAGAGGCTGACGATCAGGCGCTCATCGAC 420
421 GAAGCAATGATCAAGCTTGAATGTCACCGCAACCAAGTCCCGCTGGTGCACCAACGCAATC 480
421 GAAGCAATGATCAAGCTTGAATGTCACCGCAACCAAGTCCCGCTGGTGCACCAACGCAATC 480
481 CTGCGTGTTCATGCTGTTGCAAGGCTGCTGCTGATTCGCGAGGCTCCCACTGTTTC 540
481 CTGCGTGTTCATGCTGTTGCAAGGCTGCTGCTGATTCGCGAGGCTCCCACTGTTTC 540
541 CGCTACATCGGTGGACCAACGCAACGCTTTCAGTTTCCAGTTTCCAATGATGAACATCAAC 600
541 CGCTACATCGGTGGACCAACGCAACGCTTTCAGTTTCCAATGATGAACATCAAC 600
601 GTGGCGCTCAGCTGACCTCCGCTGTTGACGTTTCAAGGATTCATGATCGCTCCATCGGT 660
601 GTGGCGCTCAGCTGACCTCCGCTGTTGACGTTTCAAGGATTCATGATCGCTCCATCGGT 660
661 GCAGAGACCTTCTCTGAGGCTCTCCGCAACGGCGGAGGTCTACCAACGCTGAAGTCC 720
661 GCAGAGACCTTCTCTGAGGCTCTCCGCAACGGCGGAGGTCTACCAACGCTGAAGTCC 720
721 GTTCATCAAGAAAGGCGCTGTCACCGGATGTCGATGAGGCGGCTTCGCTCTCTCC 780
721 GTTCATCAAGAAAGGCGCTGTCACCGGATGTCGATGAGGCGGCTTCGCTCTCTCC 780
781 GTGGCTCACCCTGAGGCTCTTGACCTTATGCTTAAGGCAATGAGAGGCTGGCTTC 840
781 GTGGCTCACCCTGAGGCTCTTGACCTTATGCTTAAGGCAATGAGAGGCTGGCTTC 840
841 ACCCAGGCAAGGACATCGCTCTGCTGAGGCTGCTCTCTGAGTTCTTCAAGGAC 900
841 ACCCAGGCAAGGACATCGCTCTGCTGAGGCTGCTCTCTGAGTTCTTCAAGGAC 900
901 GGCACCTACCTTCAAGGCTGCGCAGCACTCCGCAAGCTGAGATGCGCAACGTTTACGCT 960
901 GGCACCTACCTTCAAGGCTGCGCAGCACTCCGCAAGCTGAGATGCGCAACGTTTACGCT 960
961 GAGCTCGTGAACGCTACCAATCGTCTCCATCGAGGACCCATCGCAGGAGATGACTGG 1020
961 GAGCTCGTGAACGCTACCAATCGTCTCCATCGAGGACCCATCGCAGGAGATGACTGG 1020
1021 GAGGTTACCAACCTCAACCGCAACCTCGGCGCAAGGTTTCAGATCGTGGCGACGAC 1080
1021 GAGGTTACCAACCTCAACCGCAACCTCGGCGCAAGGTTTCAGATCGTGGCGACGAC 1080
1081 TTCTTCTGTCACCAACCTCGAGGCGCTGAGGAGGCGATCGCTAAGAGGCTGCCAATCC 1140
1081 TTCTTCTGTCACCAACCTCGAGGCGCTGAGGAGGCGATCGCTAAGAGGCTGCCAATCC 1140
1141 ATCTGTTTAAAGTGAACAGATCGGTACCTTCAACGAGACCTTCAAGGCTGTGACATG 1200
1141 ATCTGTTTAAAGTGAACAGATCGGTACCTTCAACGAGACCTTCAAGGCTGTGACATG 1200
1201 GCTACCGGCGAGGCTACACCTCGATGATGTCGACCGTTCGGTTCAGCCGAGACAC 1260
1201 GCTACCGGCGAGGCTACACCTCGATGATGTCGACCGTTCGGTTCAGCCGAGACAC 1260
1261 ACCATTGCTGACCTCGCAGTTGCACTCAACTGTGCGCAGATCAAGACTGGTCTCCAGCA 1320

1261 ACCATTGCTGACCTCGCAGTTGCACTCAACTGTGCGCAGATCAAGACTGGTCTCCAGCA 1320
1321 CGTTCGCGAGCGTTCGCAAGGTACCAACGAGCTTCTCCGATCGAGAGCTGTTGGCGAC 1380
1321 CGTTCGCGAGCGTTCGCAAGGTACCAACGAGCTTCTCCGATCGAGAGCTGTTGGCGAC 1380
1381 GCGGCGCTTACGCAAGGTTCGAGCGCATTCACCGCTTTCAGGGCTTAATAAAGGCTT 1440
1381 GCGGCGCTTACGCAAGGTTCGAGCGCATTCACCGCTTTCAGGGCTTAATAAAGGCTT 1440
1441 TTGAGCGCGGTAACCTCAAGGTTGCGGCGCTGTTGCCCTTACTGTTTACTGGTGT 1500
1441 TTGAGCGCGGTAACCTCAAGGTTGCGGCGCTGTTGCCCTTACTGTTTACTGGTGT 1500
1501 GACTATGATCGAGGATTAATGCAAGAGCAGAGAACTCATAAAGGCTTGTTCCTGTCT 1560
1501 GACTATGATCGAGGATTAATGCAAGAGCAGAGAACTCATAAAGGCTTGTTCCTGTCT 1560
1561 CAAGCAGGGAACGTCCTT 1578
1561 CAAGCAGGGAACGTCCTT 1578

RESULT 5
AP005277
LOCUS 333150 bp DNA linear BCT 08-AUG-2002
DEFINITION Corynebacterium glutamicum ATCC 13032 DNA, complete genome, section 4/10.
ACCESSION AP005277 BA000036
VERSION AP005277.1 GI:21323710
KEYWORDS
SOURCE Corynebacterium glutamicum ATCC 13032
ORGANISM Corynebacterium glutamicum ATCC 13032
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacterineae; Corynebacteriaceae; Corynebacterium.

1
Nakagawa, S.
Complete genomic sequence of Corynebacterium glutamicum ATCC 13032
Unpublished
2 (bases 1 to 333150)
Nakagawa, S.
Direct Submission
Submitted (24-MAY-2002) Satoshi Nakagawa, Kyowa Hakko Kogyo Co. Ltd., Tokyo Research Laboratories, 3-6-6, Asahi-machi, Machida, Tokyo 194-8533, Japan (E-mail:snakagawa@axanagen.com, Tel:81-44-829-3031, Fax:81-44-813-1651)
This sequence is conducted by collaboration of Kyowa Hakko Kogyo Co. Ltd. And Kitasato University.
Location/Qualifiers
1. 333150
/organism="Corynebacterium glutamicum ATCC 13032"
/mol_type="genomic DNA"
/strain="ATCC 13032"
/db_xref="taxon:196627"
/note="ATCC 13032"
complement(63..1292)
/gene="Cg10944"
complement(63..1292)
/gene="Cg10944"
/codon_start=1
/transl_table=11
/product="Hypothetical membrane protein"
/protein_id="BAB98337.1"
/db_xref="GI:21323711"

gene
CDS
1382. .2917

IEVDGLDAPVRDTGSLSGQRQRVALARALHADAEVLVLMPTSAVDSVTVETVSIAGQ IKQLRAGKTTIIVSSSPAFYNLADRVISHV"		10107..11843 /gene="Cg10953" 10107..11843 /gene="Cg10953"		10107..11843 /gene="Cg10953"					
gene	Query Match 99.9%; Score 1576.4; DB 1; Length 333150; Best Local Similarity 99.9%; Pred. No. 4.8e-289; Matches 1577; Conservative 0; Mismatches 1; Indels 0; Gaps 0;								
CDS	1 GCCTGGGGATATGGGTAGTTTTCGCCACTAATTTCACTGATTCCTCATCGAAACAAGA 60 34849 GGCCTGGGGATATGGGTAGTTTTCGCCACTAATTTCACTGATTCCTCATCGAAACAAGA 34908								
QY	61 TTCCTGTCACAAATTTGGGTGTAGACGTGATTGAAGACATTTTGATCAGTGAATAATTTCTAG 120								
DB	34909 TTCCTGTCACAAATTTGGGTGTAGACGTGATTGAAGACATTTTGATCAGTGAATAATTTCTAG 34968								
QY	121 TTAGCTCCCAAGTTGGCATAGGAGGCCACAGTGGCTGAAATCATGCACGTATTCGCTCGC 180								
DB	34969 TTAGCTCCCAAGTTGGCATAGGAGGCCACAGTGGCTGAAATCATGCACGTATTCGCTCGC 35028								
QY	181 GAAATTTCTGACTCCCGCGGTAAACCAACCGTCGAGGACAGGTTTTCTGATGACGGT 240								
DB	35029 GAAATTTCTGACTCCCGCGGTAAACCAACCGTCGAGGACAGGTTTTCTGATGACGGT 35088								
QY	241 TCCACAGGTGTCGACGTGTTTCCATCCGGCGCATCCACCGGCTCCACGAGGCTCATGAG 300								
DB	35089 TCCACAGGTGTCGACGTGTTTCCATCCGGCGCATCCACCGGCTCCACGAGGCTCATGAG 35148								
QY	301 CTCGCTGACGGTGGCGATCGCTACCTGGGCAAGGGCGTTTTTGAAGGCAGTTGAAACAGTTC 360								
DB	35149 CTCGCTGACGGTGGCGATCGCTACCTGGGCAAGGGCGTTTTTGAAGGCAGTTGAAACAGTTC 35208								
QY	361 AACGAAGAAATCGGCGACGAGCTCGCTGGCCTTAGAGGCTGACGATCAGCGCTCATCGAC 420								
DB	35209 AACGAAGAAATCGGCGACGAGCTCGCTGGCCTTAGAGGCTGACGATCAGCGCTCATCGAC 35268								
QY	421 GAAGCAATGATCAAGCTTGAATGGCACCGCCAAACAGTCCCGCTGGGTGCAACACGCAATC 480								
DB	35269 GAAGCAATGATCAAGCTTGAATGGCACCGCCAAACAGTCCCGCTGGGTGCAACACGCAATC 35328								
QY	481 CTTGGTGTTTTCCATGGCTGTGCAAAAGGCTGTGCTGATTCCGCGAGGCTCCCACTGTTC 540								
DB	35329 CTTGGTGTTTTCCATGGCTGTGCAAAAGGCTGTGCTGATTCCGCGAGGCTCCCACTGTTC 35388								
QY	541 CGCTACATCGGTGGACCAAAACGACACAGTTTCTTCAGTTTCCAATGATGAACATCATCAAC 600								
DB	35389 CGCTACATCGGTGGACCAAAACGACACAGTTTCTTCAGTTTCCAATGATGAACATCATCAAC 35448								
QY	601 GGTGGGCGCTCAGCTGACTCCGGTGTGACGTTTACGTTTACAGGAATTCATGATCGCTCCAATCGGT 660								
DB	35449 GGTGGGCGCTCAGCTGACTCCGGTGTGACGTTTACGTTTACAGGAATTCATGATCGCTCCAATCGGT 35508								
QY	661 GCAGAGACTTCTCTGAGGCTCTCCGCAACGCGCGCGAGGCTTACCAACGCACTGAAGTCC 720								
DB	35509 GCAGAGACTTCTCTGAGGCTCTCCGCAACGCGCGCGAGGCTTACCAACGCACTGAAGTCC 35568								
QY	721 GTCATCAAGGAAAGGCGCTGTCCACCGGACTTGGCGATGAGGGGCGCTTCGCTCTCTTCC 780								
DB	35569 GTCATCAAGGAAAGGCGCTGTCCACCGGACTTGGCGATGAGGGGCGCTTCGCTCTCTTCC 35628								
QY	781 GTGGGCTCCACCGTGAAGGCTCTTGACCTTATCGTTTAAAGCAATCGAAGGCTGGCTTC 840								
DB	35629 GTGGGCTCCACCGTGAAGGCTCTTGACCTTATCGTTTAAAGCAATCGAAGGCTGGCTTC 35688								
QY	841 ACCCCAGGCAAGGACATCGCTCTTGCTCTGGACGTTGCTTCTCTGAGTTCTTCAAGGAC 900								
DB	35689 ACCCCAGGCAAGGACATCGCTCTTGCTCTGGACGTTGCTTCTCTGAGTTCTTCAAGGAC 35748								
QY	901 GGCACCTTACCACTTCGAAGGTGGCCAGCACTCCGACGCTGAGATGGCAAACTTTACGCT 960								

Db	35749	GGCACCTACCACTTCGAAGGTGGCCAGCACTCCGAGCTGAGATGCGAAACGTTTACGCT	35808
QY	961	GAGCTCGTTGACGCGTACCCCAATCGTCTCCATCGAGGACCCCACTCGAGAAAGTACGCTG	1020
Db	35809	GAGCTCGTTGACGCGTACCCCAATCGTCTCCATCGAGGACCCCACTCGAGAAAGTACGCTG	35868
QY	1021	GAGGTTACACCAACCTCACCGCAACCAATCGCGCAAGGTTTCAGATCGTTGGCGACGAC	1080
Db	35869	GAGGTTACACCAACCTCACCGCAACCAATCGCGCAAGGTTTCAGATCGTTGGCGACGAC	35928
QY	1081	TTCTTCGTACCAACCTCGAGGCGCTGAAGAGGCGCTCGCTAAGAGGCTCGCAACTCC	1140
Db	35929	TTCTTCGTACCAACCTCGAGGCGCTGAAGAGGCGCTCGCTAAGAGGCTCGCAACTCC	35988
QY	1141	ATCCTGGTTAAGGTGAACACAGATCGGTACCTCCACGAGACCTTCGAGCTCTGCACATG	1200
Db	35989	ATCCTGGTTAAGGTGAACACAGATCGGTACCTCCACGAGACCTTCGAGCTCTGCACATG	36048
QY	1201	GCTCACCGCGAGGCTACACCTCCATGATGTCCCAACCGTTCCGGTGAGACCGAGGACAC	1260
Db	36049	GCTCACCGCGAGGCTACACCTCCATGATGTCCCAACCGTTCCGGTGAGACCGAGGACAC	36108
QY	1261	ACCATTCGTGACCTCGCAGTTGCACTCAACTGTGCGCAGATCAAGACTGGTCTCCAGCA	1320
Db	36109	ACCATTCGTGACCTCGCAGTTGCACTCAACTGTGCGCAGATCAAGACTGGTCTCCAGCA	36168
QY	1321	CGTTCCGACCGTGTGCGAAAGTACAAACAGCTTCTCCGACGACGAGCTGCTTCGCGAC	1380
Db	36169	CGTTCCGACCGTGTGCGAAAGTACAAACAGCTTCTCCGACGACGAGCTGCTTCGCGAC	36228
QY	1381	GCCGCGCTTACGCAAGTTCGACGCGCATTCACGCTTTCAGGGCTAAATAAAGCGCTT	1440
Db	36229	GCCGCGCTTACGCAAGTTCGACGCGCATTCACGCTTTCAGGGCTAAATAAAGCGCTT	36288
QY	1441	TTGACGCGCGGTAACTCAAGTTTCCGCGGCTGTTGCCCTTACTACTGTTACTGTTGT	1500
Db	36289	TTGACGCGCGGTAACTCAAGTTTCCGCGGCTGTTGCCCTTACTACTGTTACTGTTGT	36348
QY	1501	GACTATGATCCAGGATTATGCAAGAGCAAGAAACTCATAAAGGCTTGTTCCTGCTCT	1560
Db	36349	GACTATGATCCAGGATTATGCAAGAGCAAGAAACTCATAAAGGCTTGTTCCTGCTCT	36408
QY	1561	CAAGCAGGGAACGTGCTT 1578	
Db	36409	CAAGCAGGGAACGTGCTT 36426	

RESULT 6

LOCUS	BX927150	348475 bp	DNA	linear	BCT 10-JUN-2004
DEFINITION	Corynebacterium glutamicum ATCC 13032, IS fingerprint type 4-5, complete genome; segment 3/10.				
ACCESSION	BX927150 BX927147				
VERSION	BX927150.1 GI:41324904				
KEYWORDS	complete genome.				
SOURCE	Corynebacterium glutamicum ATCC 13032				
ORGANISM	Corynebacterium glutamicum ATCC 13032				
REFERENCE	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacteriaceae; Corynebacterium.				
AUTHORS	1 (bases 1 to 348475) Kalinowski, J., Bathe, B., Bartels, D., Bischoff, N., Bott, M., Burkowski, A., Dusch, N., Eggeling, L., Eikmanns, B. J., Gaigalat, L., Goessmann, A., Hartmann, M., Huthmacher, K., Kramer, R., Linke, B., McHardy, A. C., Meyer, F., Mockel, B., Pfeifferle, W., Puhler, A., Rey, D. A., Ruckert, C., Rupp, O., Sahm, H., Wendisch, V. F., Wiegand, I. and Tauch, A.				
TITLE	The complete Corynebacterium glutamicum ATCC 13032 genome sequence and its impact on the production of L-aspartate-derived amino acids and vitamins				
JOURNAL	J. Biotechnol. 104 (1-3) 5-25 (2003)				
MEDLINE	22830012				
PUBMED	12948626				
REFERENCE	2 (bases 1 to 348475)				

AUTHORS
TITLE
JOURNAL
 Kalinowski, J.
 Direct Submission
 Submitted (21-JAN-2004) Joern Kalinowski, Institut fuer
 Genomforschung, Universitaet Bielefeld, Universitaetsstrasse 25,
 33615 Bielefeld, Germany
COMMENT
 E-mail: Joern.Kalinowski@Cebitec.Uni-Bielefeld.DE
 This sequence was accomplished by collaboration between Degussa AG
 and Bielefeld University.
 join(BX927148.1..1..348071, BX927149.1..51..349887,
 BX927150.1..51..348475,
 BX927151.1..51..349459, BX927152.1..51..349799, BX927153.1..51..349584,
 BX927154.1..51..349575, BX927155.1..51..349136, BX927156.1..51..349115,
 BX927157.1..51..140057).
FEATURES
 Location/Qualifiers
 1..348475
 /organism="Corynebacterium glutamicum ATCC 13032"
 /mol_type="Genomic DNA"
 /strain="DSM 20300 = ATCC 13032"
 /db_xref="taxon:196627"
 /note="IS fingerprint type: 4-5"
 complement(381..1610)
 /locus_tag="cg0783"
 complement(381..1610)
 /locus_tag="cg0783"
 /codon_start=1
 /transl_table=11
 /product="conserved hypothetical protein"
 /protein_id="CAF19386.1"
 /db_xref="GI:41324905"
 /translation="MAEKGIIVPKLSITGEGDYTLWAPNREGNEQWQAFGLGANDDLY
 VFNPSPELLVEESDDKDLTSHPEKKNFGSDAARVVPDANAEIDLIIGAPALLADRP
 SYANVKTLRSLFAVMRLDIAITAAASPVTVFSSHSVLNSVDNRGSEHYSGPNLGWETS
 VGRAVATNNGVNDALDEAITVKEVNSLITFEDAAITRIQEAQAAVQAAAEAEQIAKEE
 AEKVDYDNSPMAAGIDPIKVISIDGRTIYTLRTYLGQPVFKGFGIEITFFNSOKSL
 LRNVHDDHDLAASWSDMLGINSGETELLVHSNYSFNGLVKIDINTSDAVDT
 QMARVLEADLPVIGVDELQEGDILLIWDGHVAMYAGGGQIIEAGDPVQLNPITRTSNIG
 ESWRELEMLIKRFSKF"
 complement(1677..2537)
 /locus_tag="cg0784"
 complement(1677..2537)
 /locus_tag="cg0784"
 /codon_start=1
 /transl_table=11
 /product="PUTATIVE CELL WALL-ASSOCIATED HYDROLASE"
 /protein_id="CAF19387.1"
 /db_xref="GI:41324906"
 /translation="WIGLISLSRLATHEPAQLPOLAIPOAPDISAALGRELNSNP
 TLLSLFATSQQHLLVTSLLIQAPLIDARHDIATTAOHLVNOAGLVMSFSMPNP
 AESFAARTLISLPLMTAEASQRUEMSRQVPIVEKLVQIQILTPEPTAEVARETA
 PVYMAAQPSNDNEVGQAAVAAKEALGTPLYMGTSYSGPDCSLGTQMAWRAGVEI
 PRIADQAVGRGVYDELQEGDILLIWDGHVAMYAGGGQIIEAGDPVQLNPITRTSNIG
 AFHGYFRPTG"
 complement(2534..2797)
 /locus_tag="cg0785"
 complement(2534..2797)
 /locus_tag="cg0785"
 /codon_start=1
 /transl_table=11
 /product="hypothetical protein predicted by
 Glimmer/Critica"
 /protein_id="CAF19388.1"
 /db_xref="GI:41324907"
 /translation="MRINLPHAKELAHLCILPTPAVPALPTDGAQFDIHOALSASL
 ATYARNLLGHTAENLGNRLATGLAEIETDDQLAHALERLT"
 2896..3531
 /gene="upp"
 /locus_tag="cg0786"
 2896..3531
 /gene="upp"
 /locus_tag="cg0786"
 /EC_number="2.4.2.9"
 /codon_start=1
 /transl_table=11
 /product="PUTATIVE URACIL PHOSPHORIBOSYLTRANSFERASE"
 /protein_id="CAF19389.1"
 /db_xref="GI:41324908"
 /translation="MDITIVNHLVASRLTLLRDRSDNAAAFRAANDLGAMLIYEAS
 RDLVEHPDTKTPVAMAGTRLKOPPIIVPIIRAGLGMIDPALSMIPDAQVGFIGLAR
 DEETHEPVYLEALPQDLSNQPVFLVDPMLATGGSLHAIRLLADRGATDITACMWS
 AQPGVDALAESGLPVLVTATIDPGLDENAYIVPGLGDAGDRLYGRNIDL"
 3625..3999
 /locus_tag="cg0787"
 3625..3999
 /locus_tag="cg0787"
 /codon_start=1
 /transl_table=11
 /product="transcriptional regulator"
 /protein_id="CAF19390.1"
 /db_xref="GI:41324909"
 /translation="MSEKRWSSYGHAFASRLKKLRTLGRFSOEELADLSGVRNTISN
 YERNENKGNADVDPOLSNITRYLQAQLDVPIALMPAGSVPVAKICVDETAADIVRWPS
 EKDPILLPADLRLTRRQPGN"
 complement(3996..5588)
 /gene="pmmB"
 /locus_tag="cg0788"
 complement(3996..5588)
 /gene="pmmB"
 /locus_tag="cg0788"
 /EC_number="5.4.2.2"
 /EC_number="5.4.2.8"
 /codon_start=1
 /transl_table=11
 /product="PHOSPHOGLUCOMUTASE/PHOSPHOMANNOMUTASE"
 /protein_id="CAF19391.1"
 /db_xref="GI:41324910"
 /translation="MDESQLSFGTAGLRAPVGPARKMNVLQVTRTTAGVASWLAER
 AALNPVPHLVPEDETGIGALYPOQGLVGVYDIARYGSHFTAAITAYSEAGGEV
 TLLPTSPPLIPMLVKNGLDAGVQITASHNGAADNGYKVLNNGRLYSELGELE
 AHNAVEDPIRVPRVTRPTADQLRKYVDEMVSIVTPDQADLLRVNBSRGNLRVYTA
 LHVGGRMANAFQAPGPHTHGVKAQYDPDPTFTVAPNPEEPGSAIELLERAKEK
 NADILPALDPADRCAGVIRTAGDGGHMLSGDEVTLLATRLVPEYSGRGPVAVATT
 VVSSQLIGIIAEDKMDYSETLITGPKNLSRAADGLDGLAFAYEAVGTCPPVDVVD
 KGGISTALPMASWAALKAQASLOOKNELYRRYGFYASSQAVTSTPRELVDHWI
 AHPQBLIGVSVTHLPKQGIALHGQVGHVIRAIRGVSGTEAKAKLYLEVQQASS
 HUEAAQLHQLDEVEQSWLSKL"
 5693..6880
 /gene="amiA"
 /locus_tag="cg0789"
 5693..6880
 /gene="amiA"
 /locus_tag="cg0789"
 /codon_start=1
 /transl_table=11
 /product="PUTATIVE N-ACYL-L-AMINO ACID AMIDOHYDROLASE"
 /protein_id="CAF19392.1"
 /db_xref="GI:41324911"
 /translation="NMEIGQVVASMMDRHHEVIKWRHHLHSHPELSHMEYRTEYLA
 SVLKDHGMHPFLPGTGLMVDIGEGDSRLAFRADIDALPFLLESTGLEFSGTATGVAH
 ACCHDYVTALALACALATIELPIGIRVFOAEVMTGATDVTAHGLDGDVDAIY
 AHHVEPKLVGRVGVZAGALTASDVIEIRVKGEGGSARPHISADVVALSKLVVDL
 PGLSRKVDPTGTVLVFGTINAGYAPNALPDGSIYSGTILRTADISTWRDMRLISEL
 VEQLTAPGTVTHLIPGVPFVINDVDVATALLASAAARDMDTQSVVQAPQSSGGEDFS
 WYLEHVPGSNARLGCWPGHKQDLHQSLVDVDERAIGVGRVRLFGSLVQOQYSRSEAF
 LNS"
 7020..8429
 /gene="lpda"
 /locus_tag="cg0790"
 7020..8429
 /gene="lpda"
 /locus_tag="cg0790"
 /EC_number="1.8.1.4"
 /codon_start=1
 /transl_table=11
 /product="DIHYDROLIPOAMIDE DEHYDROGENASE"
 /protein_id="CAF19393.1"
 /db_xref="GI:41324912"

```

/translation="MAKRIVIIGGPAGYEALALAGAKYGAETVIEDVGVGSAVTMD
CVPKSFIAGTGKIDLRADDMLNRLGKSHLEIDALNIRVKOLAKQAQSEILGOL
QSDVNMINGKDFDYNTKOTQTHYIKVTHSDGSEBTEVDLVLVATGATPRILKGA
PDGERILLTWRQVYDIEELPHLILVSGSVGTGAEFVAFELGVKVTMVASRDLILPHD
DADAADVLTALAEVRSLEKHARVSVRTEDGGVYCVTRADGREIYGHSHALMTVGSII
PNTADLGLENIGVELAPSCHKVDRVSRNIPGVYAAGDCTDLFFPLASVAAQGRIAM
YHALGEGVSPFIRLKTATVATVFTPEIAAGVITHAQVDSGEVSARVILVFLATNPRAKM
RSLRHGFVKLFCRNSGLIIGVVVAPTASLEILPIAVAVTNRLTVADLADTFVAVPS
LGSITEAARQVQHQDLG"
8826..12248
/gene="pyc"
/locus_tag="cg0791"
8826..12248
/gene="pyc"
/locus_tag="cg0791"
/EC_number="6.4.1.1"
/codon_start=1
/transl_table=11
/product="PYRUVATE CARBOXYLASE"
/protein_id="CAP19394.1"
/db_xref="GI:41324913"
/translation="MSTHTSSLTLPFKKILVANRGELIAYRAFAALETGAAVAYIPR
ERGRFHSFASAVNAIGTEGSPVKAYLDIDELIQAQKVKADALIPGYFUSENAQL
ARECAENGITFIPTEVLDLTDGSRVATAKAGLPLVAESTPSKNDEIVKGAEG
QTPYIFVKAAGGGRGMRFVAPDELRLKLAATEASREAEAFDGDVAVYVERAVINPQH
IEVQILGDHTGVVHLYERDLSLORRHQKVEIAPAOHLDPBLDRICADAVKFRSL
GYGAGTVEFLVDEKNGHYFIEMNPRIOVEHTVTEVTDLVKQMLAAGATIKEL
GLTDQIKTHGAALQCRITTEDPNNGFRPDGTITIAIRSPGGAGVLDGNAOLGGEIT
AHFDSMLVMTCRGSDDEFATAVQAORALAEFTVSGVATNIGFLRLALBEDFTSKRIA
TGFIDHPHLADPQBRQIRLDYLDVTVNKPVGVRPKDVAAPIDKLPIKIDLP
Query Match
Best Local Similarity 99.9%; Score 1576.4; DB 1; Length 348475;
Matches 1577; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GGCTGGGATAGGGTAGTATTTTCGCCCACTAAATTTCAACTGATTCCTCATCGAAAACAAGA 60
DB 338411 GGCTGGGATAGGGTAGTATTTTCGCCCACTAAATTTCAACTGATTCCTCATCGAAAACAAGA 338470
QY 61 TTCGTGCAACAATTTGGGTAGACGTGATTAAGACATTTTGATTCAGTGAATATTCATG 120
DB 338471 TTCGTGCAACAATTTGGGTAGACGTGATTAAGACATTTTGATTCAGTGAATATTCATG 338530
QY 121 TTAGCTCCCAAGTTGGCATAGGAGGCCACAGTGGCTGAATCATGCACTATTCGTCCG 180
DB 338531 TTAGCTCCCAAGTTGGCATAGGAGGCCACAGTGGCTGAATCATGCACTATTCGTCCG 338590
QY 181 GAAATTTCTCGATCTCCCGCGGTAAACCAACCGTCGAGGCGAGAGGTTTCTTGATGACGGT 240
DB 338591 GAAATTTCTCGATCTCCCGCGGTAAACCAACCGTCGAGGCGAGAGGTTTCTTGATGACGGT 338650
QY 241 TCCACGGTGTCCGAGGTGTTCCATCCGGCGCATCCACCGCGGTCCACGAGGCTCATGAG 300
DB 338651 TCCACGGTGTCCGAGGTGTTCCATCCGGCGCATCCACCGCGGTCCACGAGGCTCATGAG 338710
QY 301 CTGCGTGAAGTGGCGCATCGCTACCTTGGGCAAGGGGTTTGAAGCGATTTGAAACGTC 360
DB 338711 CTGCGTGAAGTGGCGCATCGCTACCTTGGGCAAGGGGTTTGAAGCGATTTGAAACGTC 338770
QY 361 AACGAGAAATCGGCGACGAGCTCGCTGGCCCTAGAGGCTGACGATCAGCGCTCATCGAC 420
DB 338771 AACGAGAAATCGGCGACGAGCTCGCTGGCCCTAGAGGCTGACGATCAGCGCTCATCGAC 338830
QY 421 GAAGCAATGATCAAGCTTGTATGTCACCGCCAAACAGTCCCGCTGGGTGCAACACGCAATC 480
DB 338831 GAAGCAATGATCAAGCTTGTATGTCACCGCCAAACAGTCCCGCTGGGTGCAACACGCAATC 338890
QY 481 CTTGTGTTTCCATGCTGTGTCAAAGGCTGTGTCGATTCGCGAGGCTCCCACTGTTC 540
DB 338891 CTTGTGTTTCCATGCTGTGTCAAAGGCTGTGTCGATTCGCGAGGCTCCCACTGTTC 338950
QY 541 CGCTACATCGGTGGACCAACGACGATTTCTCCAGTTCCATGATGATCAATCATCAAC 600

```

RESULT 7
AX121746

```

DB 338951 CGCTACATCGGTGGACCAACGACACGCTTCTTCCAGTTTCCAATGATGAACATCATCAAC 339010
QY 601 GGTGGCGCTCACGCTGACCTCCGGTGTGACGTTTACGATTCAGGAATTCATGATCGCTCAATCGGT 660
DB 339011 GGTGGCGCTCACGCTGACCTCCGGTGTGACGTTTACGATTCAGGAATTCATGATCGCTCAATCGGT 339070
QY 661 GCAGAGACCTTCTCTGAGGCTCTCCGCCAACGGCGCGGAGGTCTACCAACGCACTGAAGTCC 720
DB 339071 GCAGAGACCTTCTCTGAGGCTCTCCGCCAACGGCGCGGAGGTCTACCAACGCACTGAAGTCC 339130
QY 721 GTCATCAAGGAAAAGGCGCTGTCCACCGGATTTGGCGATGAGGCGGCTTCGCTCCTTCC 780
DB 339131 GTCATCAAGGAAAAGGCGCTGTCCACCGGATTTGGCGATGAGGCGGCTTCGCTCCTTCC 339190
QY 781 GTCGCGCTCCACCGGTGAGGCTCTTGACCTTATCGTTTAAAGGCAATCGAAGAGGTGCGCTTC 840
DB 339191 GTCGCGCTCCACCGGTGAGGCTCTTGACCTTATCGTTTAAAGGCAATCGAAGAGGTGCGCTTC 339250
QY 841 ACCCCAGGCAAGCAATCGCTCTTGCTCTGGAGCTTGTCTCTCTGAGTTCTTCAAGGAC 900
DB 339251 ACCCCAGGCAAGCAATCGCTCTTGCTCTGGAGCTTGTCTCTCTGAGTTCTTCAAGGAC 339310
QY 901 GGCACTTACCTCTCGAAGGTGGCCAGCACTCCGAGCTGAGATGGCAAAACGTTTACGCT 960
DB 339311 GGCACTTACCTCTCGAAGGTGGCCAGCACTCCGAGCTGAGATGGCAAAACGTTTACGCT 339370
QY 961 GAGCTCGTTGACGCGTACCCCAATCGTCTCCATCGAGGACCCACTGCGAGGAAGATGATGG 1020
DB 339371 GAGCTCGTTGACGCGTACCCCAATCGTCTCCATCGAGGACCCACTGCGAGGAAGATGATGG 339430
QY 1021 GAGGCTTACCAACCTCACCGCAACCTCGCGGACAGGTTTCAGATCGTTGGCGGACGAC 1080
DB 339431 GAGGCTTACCAACCTCACCGCAACCTCGCGGACAGGTTTCAGATCGTTGGCGGACGAC 339490
QY 1081 TTCTTCTGTCACCAACCTCGAGCGCTGAGAGGCGCATCGTAAAGAGGCTGCAACTCC 1140
DB 339491 TTCTTCTGTCACCAACCTCGAGCGCTGAGAGGCGCATCGTAAAGAGGCTGCAACTCC 339550
QY 1141 ATCTCTGTTAAGGTGAACAGATCGGTACCTTCAACGAGACCTTCGACGCTGTGACATG 1200
DB 339551 ATCTCTGTTAAGGTGAACAGATCGGTACCTTCAACGAGACCTTCGACGCTGTGACATG 339610
QY 1201 GCTCACCGGCGAGGCTACACCTTCCATGATGTCCTCCAGCTTCGCGTGAGACCGAGGACCC 1260
DB 339611 GCTCACCGGCGAGGCTACACCTTCCATGATGTCCTCCAGCTTCGCGTGAGACCGAGGACCC 339670
QY 1261 ACCATTGCTGACCTCGCAGTTGCACTCAACTGTGGCCAGATCAAGACTGGTGTCTCCAGCA 1320
DB 339671 ACCATTGCTGACCTCGCAGTTGCACTCAACTGTGGCCAGATCAAGACTGGTGTCTCCAGCA 339730
QY 1321 CGTTCCGACCGGTGTCGAAAAGTACAAACAGCTTCTCCGATTCGAGCAGCTGTTGGCGAC 1380
DB 339731 CGTTCCGACCGGTGTCGAAAAGTACAAACAGCTTCTCCGATTCGAGCAGCTGTTGGCGAC 339790
QY 1381 GCGCGCTCTACGAGGTCGAGCGCATTCGACGCTTTCAGGGCTAAATAAAGCGCTT 1440
DB 339791 GCGCGCTCTACGAGGTCGAGCGCATTCGACGCTTTCAGGGCTAAATAAAGCGCTT 339850
QY 1441 TTGACGCGCGGTAACTCAAGTTGCGGCGCTGTTGCTTACTACTGTTTACTGTTGT 1500
DB 339851 TTGACGCGCGGTAACTCAAGTTGCGGCGCTGTTGCTTACTACTGTTTACTGTTGT 339910
QY 1501 GACTATGATCGAGGATTTATGGCAAGCAGAAAGAACTCATAAAGGCGCTTGTTCCTGTCT 1560
DB 339911 GACTATGATCGAGGATTTATGGCAAGCAGAAAGAACTCATAAAGGCGCTTGTTCCTGTCT 339970
QY 1561 CAAGCAGGAAACGTGCTT 1578
DB 339971 CAAGCAGGAAACGTGCTT 339988

```

LOCUS AX127146 34980 bp DNA linear PAT 11-MAY-2001
DEFINITION Sequence 7062 from Patent EP1108790.
ACCESSION AX127146 AX114121
VERSION AX127146.1 GI:14041134
KEYWORDS Corynebacterium glutamicum
SOURCE Corynebacterium glutamicum
ORGANISM Corynebacterium glutamicum
REFERENCE 1
AUTHORS Nakagawa, S., Mizoguchi, H., Ando, S., Hayashi, M., Ochiai, K., Yokoi, H., Tateishi, N., Senoh, A., Ikeda, M. and Ozaki, A.
TITLE Novel polynucleotides
JOURNAL Patent: EP 1108790-A 7062 20-JUN-2001;
KYOWA HAKKO KOGYO CO., LTD. (JP)
FEATURES
source 1..34980
/organism="Corynebacterium glutamicum"
/mol_type="genomic DNA"
/db_xref="taxon:1718"
/notes="Seq 1 to long (3.309.400) split in 11, seq 7062
0.900.001 1.249.980"
ORIGIN
Query Match 99.9%; Score 1576.4; DB 6; Length 349980;
Best Local Similarity 99.9%; Pred. No. 4.8e-289;
Matches 1577; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GGCTGGGATATGGTAGTATTTCCGCACTAATTTCAACTATGTCCTCATCGAACAAGA 60
DB 134799 GGCTGGGATATGGTAGTATTTCCGCACTAATTTCAACTATGTCCTCATCGAACAAGA 134858
QY 61 TTCTGCAACAATTTGGGTAGACGTGATTGAAGACATTTGATCAGCTGAATATTTAG 120
DB 134859 TTCTGCAACAATTTGGGTAGACGTGATTGAAGACATTTGATCAGCTGAATATTTAG 134918
QY 121 TTAGTCTCCCAAGTTGGCATAGGAGGCCACAGTGGCTGAATATCATGCAATTCGCTCGC 180
DB 134919 TTAGTCTCCCAAGTTGGCATAGGAGGCCACAGTGGCTGAATATCATGCAATTCGCTCGC 134978
QY 181 GAAATTTCTCACTCCCGGTGTAACCAACCGTCGAGGACAGGTTTTCCTGGATGAGGT 240
DB 134979 GAAATTTCTCACTCCCGGTGTAACCAACCGTCGAGGACAGGTTTTCCTGGATGAGGT 135038
QY 241 TCCACCGGTTCGACAGGTGTTTCATCCGGCGCATCCACCGCGTCACACGAGGCTCATGAG 300
DB 135039 TCCACCGGTTCGACAGGTGTTTCATCCGGCGCATCCACCGCGTCACACGAGGCTCATGAG 135098
QY 301 CTGCTGACCGGTGGCGATCGTACCTGGGCAAGGGCGTTTGAAGGCGATTTGAAGCGTC 360
DB 135099 CTGCTGACCGGTGGCGATCGTACCTGGGCAAGGGCGTTTGAAGGCGATTTGAAGCGTC 135158
QY 361 AAGCAAGAAATCGGCGACGAGCTCGTGGCTAGAGGCTGACGATCAGCGCTCATCGAC 420
DB 135159 AAGCAAGAAATCGGCGACGAGCTCGTGGCTAGAGGCTGACGATCAGCGCTCATCGAC 135218
QY 421 GAAGCAATGATCAAGCTTGAATGGCAGCCGCAACAAAGTCCCGCTGGGTGCAAGCGCAATC 480
DB 135219 GAAGCAATGATCAAGCTTGAATGGCAGCCGCAACAAAGTCCCGCTGGGTGCAAGCGCAATC 135278
QY 481 CTTGGTGTTCATGCTGTGTTGCAAGGCTGCTGCTGATTTCCGAGCGCTCCCACTGTTTC 540
DB 135279 CTTGGTGTTCATGCTGTGTTGCAAGGCTGCTGCTGATTTCCGAGCGCTCCCACTGTTTC 135338
QY 541 CGCTACATCGGTGGAACAAACGACACAGTTCCTTCAGTTTCGATGATGATGATGATGATCAAC 600
DB 135339 CGCTACATCGGTGGAACAAACGACACAGTTCCTTCAGTTTCGATGATGATGATGATGATCAAC 135398
QY 601 GGTGGGCTCAGCTGACTCGGTGTTGAGCTTCAGGAATTCATGATCGCTCCCAATCGGT 660
DB 135399 GGTGGGCTCAGCTGACTCGGTGTTGAGCTTCAGGAATTCATGATCGCTCCCAATCGGT 135458
QY 661 GCAGAGACCTTCTCTGAGGCTCTCCGCAACGGCGCGGAGGTCTACCAACGCACTGAAAGTCC 720

DB 135459 GCAGAGACCTTCTCTGAGGCTCTCCGCAACGGCGGAGGTCTACCAACGCACTGAAAGTCC 135518
QY 721 GTCATCAAGGAAAAGGGCTCTGTCACCGGACTTTGGCGATTTGGCGGCTTCGCTCTCTTC 780
DB 135519 GTCATCAAGGAAAAGGGCTCTGTCACCGGACTTTGGCGATTTGGCGGCTTCGCTCTCTTC 135578
QY 781 GTGGGCTTCCACCGGTGAGGCTCTTGTACCTTATCGTTTAAAGCAATCAGAAAGGCTTTC 840
DB 135579 GTGGGCTTCCACCGGTGAGGCTCTTGTACCTTATCGTTTAAAGCAATCAGAAAGGCTTTC 135638
QY 841 ACCCGAGGCAAGGACATCGCTCTTGTCTGAGGCTTGTCTCTCTGAGTCTTTCAGAGGAC 900
DB 135639 ACCCGAGGCAAGGACATCGCTCTTGTCTGAGGCTTGTCTCTCTGAGTCTTTCAGAGGAC 135698
QY 901 GGCACCTTACCACCTTCAAGGTGGCGAGCACTCCGAGCTGAGATGCAAAACGTTTACGCT 960
DB 135699 GGCACCTTACCACCTTCAAGGTGGCGAGCACTCCGAGCTGAGATGCAAAACGTTTACGCT 135758
QY 961 GAGCTCGTTGAGCGCTTACCAATTCGTTCTCATCGAGGACCCACTCGAGGAAGTATGATGG 1020
DB 135759 GAGCTCGTTGAGCGCTTACCAATTCGTTCTCATCGAGGACCCACTCGAGGAAGTATGATGG 135818
QY 1021 GAGGTTTACCAACCTTACCGCAACCATCGGCAACAGTTTCAGATCGTTGGCGAGGAC 1080
DB 135819 GAGGTTTACCAACCTTACCGCAACCATCGGCAACAGTTTCAGATCGTTGGCGAGGAC 135878
QY 1081 TTCTTCTGTCACCAACCTTGAAGGAGGCGCATCGTAAAGAGGCTGCAACTCC 1140
DB 135879 TTCTTCTGTCACCAACCTTGAAGGAGGCGCATCGTAAAGAGGCTGCAACTCC 135938
QY 1141 ATCTCTGTTTAAAGTGAAACAGATCGGTACCTTCCACGAGACCTTTCGACGCTGTCGACATG 1200
DB 135939 ATCTCTGTTTAAAGTGAAACAGATCGGTACCTTCCACGAGACCTTTCGACGCTGTCGACATG 135998
QY 1201 GCTCACCGGCGAGGCTTACACCTTCATGATGTCCACCGTTCGCGTGAGACCGAGGACACC 1260
DB 135999 GCTCACCGGCGAGGCTTACACCTTCATGATGTCCACCGTTCGCGTGAGACCGAGGACACC 136058
QY 1261 ACCATTGCTGACCTCGCAGTTGCACTCAACTGTGGCGCAGATCAAGACTTGTGCTCCAGCA 1320
DB 136059 ACCATTGCTGACCTCGCAGTTGCACTCAACTGTGGCGCAGATCAAGACTTGTGCTCCAGCA 136118
QY 1321 CGTTCGACCGGTTCGCAAAAGTACAAACAGCTTCTCCGATCGAGCAGCTGCTTGGCGAC 1380
DB 136119 CGTTCGACCGGTTCGCAAAAGTACAAACAGCTTCTCCGATCGAGCAGCTGCTTGGCGAC 136178
QY 1381 GCGGCGTCTACGACGCTGCGAGCGCATTTCCACGCTTTCAGGGCTTAAATAAAAGCGCTT 1440
DB 136179 GCGGCGTCTACGACGCTGCGAGCGCATTTCCACGCTTTCAGGGCTTAAATAAAAGCGCTT 136238
QY 1441 TTCGACCGCGGTAACCTCAAGTTTCGCGGGCGTGTGCTTACTTACTGTTACTGTTGTTGT 1500
DB 136239 TTCGACCGCGGTAACCTCAAGTTTCGCGGGCGTGTGCTTACTTACTGTTACTGTTGTTGT 136298
QY 1501 GACTATGATCGAGGATTTATGGCAAAAGCAGAGAAAATCTCATAAAGGCTTGTTCCTGCT 1560
DB 136299 GACTATGATCGAGGATTTATGGCAAAAGCAGAGAAAATCTCATAAAGGCTTGTTCCTGCT 136358
QY 1561 CAAGCAGGGAACGTGCTT 1578
DB 136359 CAAGCAGGGAACGTGCTT 136376

RESULT 8
AX763190
LOCUS
DEFINITION Sequence 7 from Patent WO03040291.
ACCESSION AX763190
VERSION AX763190.1 GI:32257758
KEYWORDS Corynebacterium glutamicum
SOURCE Corynebacterium glutamicum
ORGANISM

AX763190 1405 bp DNA linear PAT 25-JUN-2003
Sequence 7 from Patent WO03040291.

Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Corynebacteriaceae; Corynebacterium.

1
Zelder, O., Pompejus, M., Schroeder, H., Kroeger, B., Kloppe, C. and
Haberkauer, G.
Genes encoding for carbon metabolism and energy-producing proteins
Patent: WO 03040291-A 7 15-MAY-2003;
BASP AKTIENGESellschaft (DE)

FEATURES

source

1. 1405
/organism="Corynebacterium glutamicum"
/mol_type="unassigned DNA"
/db_xref="taxon:1718"
101. 1378
/note="unnamed protein product; RXA00235"
/codon_start=1
/transl_table=11
/protein_id="CAD99859.1"
/db_xref="GI:32257759"
/translation="MAEIMHVPAREILDRSGNPTVEAEVFLDQSHGVGVPSTG
VHEAHLGGDRYKGLKAVENNEBIDELAGLEADQRLIDRAMIKLDGTANK
SRLGANALIGVSMVAKAADSAGLPLFRYIGPNARHLVPMNMNITGGHADSVD
VQEFMIAPGAEPTSEALRNGEVHAKSVIKELSTGLGDEGGPAPSVGSTRAL
DLIVAEKAGTPGKDLALDVAASSFFKDDTHFEGGQSHSAEMANVYAEVLVDAY
PIVSTEDLQDDMEGYTNLTATIGDKQIVGDFFVTNPERLKEGIARAKAASILVK
VNQIGLTFETDAVMAHRAGYTSNMHSRGETEDTTLADLVALNCGIINTGAPRS
DRVAKYNQLLRLEQLDGVYVAGRSAPRFG"

CDS

ORIGIN

Query Match 88.8%; Score 1401.8; DB 6; Length 1405;
Best Local Similarity 99.9%; Pred. No. 9.1e-256;
Matches 1403; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

51 CGAAACAAGATTCGTCGAACAATTCGGGTAGACGTGATTGAAGACATTTGATCAGTGA 110
1 CGAAACAAGATTCGTCGAACAATTCGGGTAGACGTGATTGAAGACATTTGATCAGTGA 60

111 ATAATCTAGTTAGTCCCAAGTTCGCATAGAGGACCAAGTGGTGAATCATCAGT 170
61 ATAATCTAGTTAGTCCCAAGTTCGCATAGAGGACCAAGTGGTGAATCATCAGT 120

171 ATTCGCTCGGAAATTCGATCCCGCGGTAAACCAACCGTCGAGGACAGAGTTTCT 230
121 ATTCGCTCGGAAATTCGATCCCGCGGTAAACCAACCGTCGAGGACAGAGTTTCT 180

231 GGATGACGGTCCCAAGTTCGAGTTCGAGTTCCTCCGCGGATCCACCGGCGTCCAGA 290
181 GGATGACGGTCCCAAGTTCGAGTTCGAGTTCCTCATCCGCGGATCCACCGGCGTCCAGA 240

291 GGCTCATGAGTTCGAGTTCGAGTTCGATCCGTCGATCCGTCGAGGACGGCGTTTGAAGGCA 350
241 GGCTCATGAGTTCGAGTTCGAGTTCGATCCGTCGATCCGTCGAGGACGGCGTTTGAAGGCA 300

351 TGAACACGTCAACGAAGAAATCGGCGACGAGTTCGTCGCTGAGGCTGACGATCAGCG 410
301 TGAACACGTCAACGAAGAAATCGGCGACGAGTTCGTCGCTGAGGCTGACGATCAGCG 360

411 CCTCATGACGAAGCAATGATCAAGTTCGATCGGACCGGCAACAAAGTCCCGCTGGGTGC 470
361 CCTCATGACGAAGCAATGATCAAGTTCGATCGGACCGGCAACAAAGTCCCGCTGGGTGC 420

471 AAACGCAATCTTGGTGTTCATGGCTGTTCAAGGCTGCTGATCCGACGGGCT 530
421 AAACGCAATCTTGGTGTTCATGGCTGTTCAAGGCTGCTGATCCGACGGGCT 480

531 CCCAGTTCGCTACATCGGTGGACCAACCAACGATTCCTCCAGTTCGAATGATGAA 590
481 CCCAGTTCGCTACATCGGTGGACCAACCAACGATTCCTCCAGTTCGAATGATGAA 540

591 CATCATCAACGGTGGGCTCAGCTGATCCGGTGTTCAGCTTCAGGATTCATGATCGC 650
541 CATCATCAACGGTGGGCTCAGCTGATCCGGTGTTCAGCTTCAGGATTCATGATCGC 600

QY 651 TCCAAATCGGTGACAGAGACTTCTCTGAGGCTTCTCGCAACCGCGGAGGCTTACCAAGC 710
DB 601 TCCAAATCGGTGACAGAGACTTCTCTGAGGCTTCTCGCAACCGCGGAGGCTTACCAAGC 660

QY 711 ACTGAAGTTCGTCATCAAGGAAAAGGGCTGTCACCGGACTTGGCGATGAGGGCGGCTT 770
DB 661 ACTGAAGTTCGTCATCAAGGAAAAGGGCTGTCACCGGACTTGGCGATGAGGGCGGCTT 720

QY 771 CGCTCCTTTCGCTCGGCTCCACCGGTGAGGCTTGTGACCTTATCGTTAAGGCAATCGAGA 830
DB 721 CGCTCCTTTCGCTCGGCTCCACCGGTGAGGCTTGTGACCTTATCGTTAAGGCAATCGAGA 780

QY 831 GGCTGGCTTCAACCGGAGGAGGAGTCTGCTTGTGAGGCTTGTGCTTCTCTGAGT 890
DB 781 GGCTGGCTTCAACCGGAGGAGGAGTCTGCTTGTGAGGCTTGTGCTTCTCTGAGT 840

QY 891 CTTCAAGGACGGCACTTACCACTTGAAGTGGGCGGAGTCTCCGAGCTGAGTGGCAAA 950
DB 841 CTTCAAGGACGGCACTTACCACTTGAAGTGGGCGGAGTCTCCGAGCTGAGTGGCAAA 900

QY 951 GGTTCAGCTGAGTCTGTTGACGGTACCAATCGTCTTCAATCGAGGACCCACTGCAGCA 1010
DB 901 GGTTCAGCTGAGTCTGTTGACGGTACCAATCGTCTTCAATCGAGGACCCACTGCAGCA 960

QY 1011 AGATGACTGGGAGGTTTACCAACCTCACCGCAACCATCGCGACAGGTTTCAGATCGT 1070
DB 961 AGATGACTGGGAGGTTTACCAACCTCACCGCAACCATCGCGACAGGTTTCAGATCGT 1020

QY 1071 TGGCGAGCACTTTCGTCACCAACCTGAGGCGCTGAGGAGGCGATCGCTAAGAGGC 1130
DB 1021 TGGCGAGCACTTTCGTCACCAACCTGAGGCGCTGAGGAGGCGATCGCTAAGAGGC 1080

QY 1131 TGGCAACTCCATCTCTGTTAAGGTGAACCGTACCGTACCGTACCGAGCTTCGAGC 1190
DB 1081 TGGCAACTCCATCTCTGTTAAGGTGAACCGTACCGTACCGTACCGAGCTTCGAGC 1140

QY 1191 TGTGCAATGGCTCACCGCGGAGGCTACACCTCCATGATGTCACCGTTCGGGTGAGC 1250
DB 1141 TGTGCAATGGCTCACCGCGGAGGCTACACCTCCATGATGTCACCGTTCGGGTGAGC 1200

QY 1251 CGAGGACACCACTTCGACCTCGAGTTCGACTCAACTGTGGCGGAGATCAAGACTGG 1310
DB 1201 CGAGGACACCACTTCGACCTCGAGTTCGACTCAACTGTGGCGGAGATCAAGACTGG 1260

QY 1311 TGCTCCGAGGCTTCGACCGGCTGTCGAAAGTACAAAGTTCCTCCGATCGAGGCT 1370
DB 1261 TGCTCCGAGGCTTCGACCGGCTGTCGAAAGTACAAAGTTCCTCCGATCGAGGCT 1320

QY 1371 GCTTGGCGAGCGCGGCTTACGAGGCTCGAGGCGCATTCGACGCTTCAGGGCTAAAT 1430
DB 1321 GCTTGGCGAGCGCGGCTTACGAGGCTCGAGGCGCATTCGACGCTTCAGGGCTAAAT 1380

QY 1431 AAAAGCGCTTTTCGACCGCGGTAA 1455
DB 1381 AAAAGCGCTTTTCGACCGCGGTAA 1405

RESULT 9
AX064945
LOCUS
DEFINITION Sequence 71 from Patent WO0100844.
ACCESSION AX064945
VERSION AX064945.1 GI:12542657
KEYWORDS
SOURCE
ORGANISM
Corynebacterium glutamicum
Corynebacterium glutamicum
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Corynebacteriaceae; Corynebacterium.
1
Pompejus, M., Kroeger, B., Schroeder, H., Zelder, O. and Haberkauer, G.
corynebacterium glutamicum genes encoding proteins involved in
carbon metabolism and energy production
Patent: WO 0100844-A 71 04 JAN-2001;
JOURNAL

BASF AKTIENGESELLSCHAFT (DE)
 Location/Qualifiers
 1..1398
 /organism="Corynebacterium glutamicum"
 /mol_type="unassigned DNA"
 /db_xref="taxon:1718"
 101..1378
 /note="unassigned protein product; RXA00235"
 /codon_start=1
 /transl_table=11
 /protein_id="CAC25712.1"
 /db_xref="GI:12542658"
 /translation="MAEIMHVPAREILDSRNPVEAEVFLDDSGHVGAVPVSASTG
 VHEAHELDDGGRIYKGLVKAENVNEEIGDELAGEADQRLIDRAMIKLDGTANK
 SRLGNATLGVSMVAKAADAGSLPLFRYIGFPAHVLVPMNNITGGHADSQVD
 VQEFMIAPIGATFSEALRNGAEVYHALKSVIGKGLSTGLGDEGGAPSVGSVREAL
 DLIVBAIEKQGTPOKDALALDVASBPFDGTVHPFGGSHSAEMANVYARILVAY
 PIVSTEDPQDDDEGVYTNLATIGDKVQIVGDDPFTVTPERLEKGIKKAANSILVK
 VNQIGTLTETDAVDMHRAGYTSMHSRSETEDTIIADLAVALNCGQIKTGPAPRS
 DRVAKYNOLLRIEQLLDAGVYGRSAPFRPQG"

ORIGIN

Query Match 88.4%; Score 1394.8; DB 6; Length 1398;
 Best Local Similarity 99.9%; Pred. No. 1.9e-254;
 Matches 1396; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

51 CGAAACAAGATTCGTCACCAATTTGGGTGTAGACGTGATTAAGACATTTGATCACGTGA 110
 1 CGAAACAAGATTCGTCACCAATTTGGGTGTAGACGTGATTAAGACATTTGATCACGTGA 60
 111 ATATTCTAGTTAGTCCCAAGTTGGCATAGGAGCCACAGTGGTGAATCATGACAGT 170
 61 ATAATTCAGTTAGTCCCAAGTTGGCATAGGAGCCACAGTGGTGAATCATGACAGT 120
 171 ATTGCTCGCAAAATTCGACTCCCGCGGTAAACCCACCGTCCGAGGCGAGAGTTTCT 230
 121 ATTGCTCGCAAAATTCGACTCCCGCGGTAAACCCACCGTCCGAGGCGAGAGTTTCT 180
 231 GGATGACGGTTCCCAAGTGTGCGAGGTGTTCCATCCGCGGCATCCACCGCGTCCACGA 290
 181 GGATGACGGTTCCCAAGTGTGCGAGGTGTTCCATCCGCGGCATCCACCGCGTCCACGA 240
 291 GGCTCATGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 350
 241 GGCTCATGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 300
 351 TGAATAACGTCAACGAAGAAATCGCGACGAGTCTGCTGGCCCTAGAGGCTGACGATCAGCG 410
 301 TGAATAACGTCAACGAAGAAATCGCGACGAGTCTGCTGGCCCTAGAGGCTGACGATCAGCG 360
 411 CCTCATCGAAGAAAGATGATCAAGTGTGATGGCAACCGCCAAAGTCCCGCTGGGTGC 470
 361 CCTCATCGAAGAAAGATGATCAAGTGTGATGGCAACCGCCAAAGTCCCGCTGGGTGC 420
 471 AAAGCAATCTTGGTGTTCATGGCTGTGTTGCAAGGCTGCTGCTGATTCGCGAGGCT 530
 421 AAAGCAATCTTGGTGTTCATGGCTGTGTTGCAAGGCTGCTGCTGATTCGCGAGGCT 480
 531 CCCACTGTTCCGCTACATCGGTGGACCAAGACGACAGTCTTCCAGTTCCTCAATGATGA 590
 481 CCCACTGTTCCGCTACATCGGTGGACCAAGACGACAGTCTTCCAGTTCCTCAATGATGA 540
 591 CATCATCAAGGTTGGGCTGACGTTGATCGGTGTTGACGTTTCAAGAAATCATGATCGC 650
 541 CATCATCAAGGTTGGGCTGACGTTGATCGGTGTTGACGTTTCAAGAAATCATGATCGC 600
 651 TCCAAATCGGTGACAGACCTTCTCTGAGGCTCTCGCAACCGCGCGAGGCTTACCAAGC 710
 601 TCCAAATCGGTGACAGACCTTCTCTGAGGCTCTCTCGCAACCGCGCGAGGCTTACCAAGC 660
 711 ACTGAAGTCCGTATCAAGGAAAGAGGCTCTCTCAACCGGACTTTCGCGATGAGGCGGCTT 770
 661 ACTGAAGTCCGTATCAAGGAAAGAGGCTCTCTCAACCGGACTTTCGCGATGAGGCGGCTT 720

Qy 771 CGCTCTCTTCGCTGGCTCCACCCGTTGAGGCTCTTGACCTTATCGTTAAGCAATCGAGAA 830
 Db 721 CGCTCTCTTCGCTGGCTCCACCCGTTGAGGCTCTTGACCTTATCGTTAAGCAATCGAGAA 780
 Qy 831 GGCTGGCTTCACCCGAGGCAAGACATCGCTCTTGCTCTGCGAGTGGCTTCTCTGAGTT 890
 Db 781 GGCTGGCTTCACCCGAGGCAAGACATCGCTCTTGCTCTGCGAGTGGCTTCTCTGAGTT 840
 Qy 891 CTTCAAGGAGCGGACCTTACCACTTCGAAGGTGGCCAGCACTCCGACGCTGAGATGGCAAA 950
 Db 841 CTTCAAGGAGCGGACCTTACCACTTCGAAGGTGGCCAGCACTCCGACGCTGAGATGGCAAA 900
 Qy 951 CGTTTACGCTGAGCTCGTTGACCGGTACCAATCGTCTCCATCGAGGACCACTTCGACGGA 1010
 Db 901 CGTTTACGCTGAGCTCGTTGACCGGTACCAATCGTCTCCATCGAGGACCACTTCGACGGA 960
 Qy 1011 AGATGACTGGAGGGTTTACACCACTCAGCAACCTCAGCAACCTCAGCAACCTCAGCAACCT 1070
 Db 961 AGATGACTGGAGGGTTTACACCACTCAGCAACCTCAGCAACCTCAGCAACCTCAGCAACCT 1020
 Qy 1071 TGGCGACGACTTCTTCTGCTACCAACCTCAGCGCTGAGGCGCATCGCTAAGAAAGGC 1130
 Db 1021 TGGCGACGACTTCTTCTGCTACCAACCTCAGCGCTGAGGCGCATCGCTAAGAAAGGC 1080
 Qy 1131 TGGCAACTCCATCCTGTTAAGGTGAACCAAGATCGGTACCCCTCACCGAGACCTTTCGACGC 1190
 Db 1081 TGGCAACTCCATCCTGTTAAGGTGAACCAAGATCGGTACCCCTCACCGAGACCTTTCGACGC 1140
 Qy 1191 TGTGACATGGCTCAGCGCGAGGCTACACCTCCATGATGTCCACCGTTCGGTGAAGC 1250
 Db 1141 TGTGACATGGCTCAGCGCGAGGCTACACCTCCATGATGTCCACCGTTCGGTGAAGC 1200
 Qy 1251 CGAGGACACCACTTGTGACCTCGCAGTTGCACTCAACTGTGGCCAGATCAAGACTGG 1310
 Db 1201 CGAGGACACCACTTGTGACCTCGCAGTTGCACTCAACTGTGGCCAGATCAAGACTGG 1260
 Qy 1311 TGCTCAGACGCTTCCGACCGGTGTCGCAAGTGAACCAAGTGAACCAAGTGAACCAAGTGA 1370
 Db 1261 TGCTCAGACGCTTCCGACCGGTGTCGCAAGTGAACCAAGTGAACCAAGTGAACCAAGTGA 1320
 Qy 1371 GCTTGGCGACGCGGGCTTACGAGGTGCGAGGCAATCCCAAGCTTTCAGGGCTTAAT 1430
 Db 1321 GCTTGGCGACGCGGGCTTACGAGGTGCGAGGCAATCCCAAGCTTTCAGGGCTTAAT 1380
 Qy 1431 AAAAGCGCTTTTCGAGCG 1448
 Db 1381 AAAAGCGCTTTTCGAGCG 1398

RESULT 10
 BD163286
 LOCUS BD163286
 DEFINITION Novel polynucleotide.
 ACCESSION BD163286
 VERSION BD163286.1 GI:27869050
 KEYWORDS JP 2002191370-A/1085.
 SOURCE unidentified
 ORGANISM unclassified.
 REFERENCE 1 (bases 1 to 1275)
 AUTHORS Nakagawa, S., Mizoguchi, H., Ando, S., Hayashi, M., Ochiai, K.,
 Yokoi, H., Tateishi, N., Senoo, A., Ikeda, M. and Ozaki, A.
 TITLE Novel polynucleotide
 JOURNAL Patent: JP 2002191370-A 1085 09-JUL-2002;
 COMMENT KYOWA HAKKO KOGYO CO LTD
 OS Corynebacterium glutamicum
 PN JP 2002191370-A/1085
 PD 09-JUL-2002
 PF 15-DEC-2000 JP 2000405096
 PI SATOSHI NAKAGAWA, HIROSHI MIZOGUCHI, SEIKO ANDO, MIKIO HAYASHI,
 KEIKO OCHIAI,
 PI HARUHIKO YOKOI, NAKO TATEISHI, AKIHIRO SENOO, MASATO IKEDA, AKIO

```

PI OZAKI
PC C12N15/09, C12N15/09, C07K14/34, C07K16/12, C07K16/40, C12M1/00, PC
C12N1/15,
PC C12N1/19, C12N1/21, C12N5/10, C12N9/00, C12N9/02, C12P7/40, C12P13/
PC 04, C12P13/08,
PC C12P19/00, C12P19/34, C12P21/02, C12Q1/37, C12Q1/68, G01N33/53, PC
G01N33/566,
PC G01N33/569, G01N33/68, G01N37/00//C12P21/08, (C12N1/21, C12R1:15),
PC (C12N1/21, C12R1:13), (C12N1/21, C12R1:01), (C12P13/08, C12R1:15),
PC C12N15/00,
PC C12N5/00, C12N15/00
CC Novel polynucleotide
FH Key location/Qualifiers
FT source 1..1275
FT /organism='Corynebacterium glutamicum'.
FEATURES
    source
        Location/Qualifiers
            1..1275
                /organism='unidentified'
                /mol_type='genomic DNA'
                /db_xref='taxon:32644'
ORIGIN
    Query Match 80.7%; Score 1273.4; DB 6; Length 1275;
    Best Local Similarity 99.9%; Pred. No. 2.1e-231;
    Matches 1274; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 151 GTGGCTGAATCATGACGCTATTGCTCGCGAAATCTCGACTCCCGCGGTAAACCAACC 210
DB 1 GTGGCTGAATCATGACGCTATTGCTCGCGAAATCTCGACTCCCGCGGTAAACCAACC 60
QY 211 GTGAGGCGAGAGGTTTCTGATGACGTTTCCACGGTGTCCAGAGTTCGCAAGTCCCGGC 270
DB 61 GTGAGGCGAGAGGTTTCTGATGACGTTTCCACGGTGTCCAGAGTTCGCAAGTCCCGGC 120
QY 271 GCATCCACCGGCTCCACAGGCTCATGAGCTGCGGTGACGGTGGCGATCGCTACCTGGGC 330
DB 121 GCATCCACCGGCTCCACAGGCTCATGAGCTGCGGTGACGGTGGCGATCGCTACCTGGGC 180
QY 331 AAGGGCGTTTGAAGGCGAGTGAACAGTCAACGAAGAAATCGGCGACGAGCTCGCTGGC 390
DB 181 AAGGGCGTTTGAAGGCGAGTGAACAGTCAACGAAGAAATCGGCGACGAGCTCGCTGGC 240
QY 391 CTAGAGGCTGACGATCAGCGCTCATCGACGAGCAATGATCAAGTTGATGGCACCGCC 450
DB 241 CTAGAGGCTGACGATCAGCGCTCATCGACGAGCAATGATCAAGTTGATGGCACCGCC 300
QY 451 AACAGTCCCGCTGGTGTCAACGCAATCTTGTGTTTCCATGGCTGTGCAAGGCT 510
DB 301 AACAGTCCCGCTGGTGTCAACGCAATCTTGTGTTTCCATGGCTGTGCAAGGCT 360
QY 511 GCTGCTGATTCGCGAGGCTCCACCTGTTCCGCTACATCGGTGGACCAACGACACGTT 570
DB 361 GCTGCTGATTCGCGAGGCTCCACCTGTTCCGCTACATCGGTGGACCAACGACACGTT 420
QY 571 CTTCCAGTTCCCAATGATGAACATCATCAACGGTGGCGCTCAGCTGACTCCGGTGTGAC 630
DB 421 CTTCCAGTTCCCAATGATGAACATCATCAACGGTGGCGCTCAGCTGACTCCGGTGTGAC 480
QY 631 GTTCAGGAATTCATGATCGCTCAATCGGTGAGAGACCTTCTGTGAGGCTTCGCGAAC 690
DB 481 GTTCAGGAATTCATGATCGCTCAATCGGTGAGAGACCTTCTGTGAGGCTTCGCGAAC 540
QY 691 GCGCGGAGGCTTACACGACGAGTCCGCTCATCAGGAAGGCGTGTCCACCGGA 750
DB 541 GCGCGGAGGCTTACACGACGAGTCCGCTCATCAGGAAGGCGTGTCCACCGGA 600
QY 751 CTTGCGATGAGGGGCGCTTCGCTCCTTCGCTCGGCTCCACCGGTGAGGCTCTTGACCTT 810
DB 601 CTTGCGATGAGGGGCGCTTCGCTCCTTCGCTCGGCTCCACCGGTGAGGCTCTTGACCTT 660
QY 811 ATCGTTAAGGCAATCGAAGGCTGGCTTTCACCCAGGCAAGGACATCGCTTGTGCTGTG 870
DB 661 ATCGTTAAGGCAATCGAAGGCTGGCTTTCACCCAGGCAAGGACATCGCTTGTGCTGTG 720

```

```

QY 871 GACGTTGCTTCTCTGAGTTCTTCAAGGACGCACTTACCACTTGAAGGTGGCCAGCAC 930
DB 721 GACGTTGCTTCTCTGAGTTCTTCAAGGACGCACTTACCACTTGAAGGTGGCCAGCAC 780
QY 931 TCCGAGCTGAGATGGCAAAAGTTTACGCTGAGCTCGTTGACGCGTACCAATCGTCTCC 990
DB 781 TCCGAGCTGAGATGGCAAAAGTTTACGCTGAGCTCGTTGACGCGTACCAATCGTCTCC 840
QY 991 ATCGAGGACCACTTCAGGAGATGACTGGGAGGTTTACCAAACTTACCCGCAACCATC 1050
DB 841 ATCGAGGACCACTTCAGGAGATGACTGGGAGGTTTACCAAACTTACCCGCAACCATC 900
QY 1051 GCGCAAAAGTTTCAGATCGTTGGCGACGACTTCTTCTGTCACCAACCTTGAGCGCTGAAG 1110
DB 901 GCGCAAAAGTTTCAGATCGTTGGCGACGACTTCTTCTGTCACCAACCTTGAGCGCTGAAG 960
QY 1111 GAGGCGATCGTTAAGAGGCTGCCAACTTCCATCTCTGTTAAGGTGAACAGATCGGTACC 1170
DB 961 GAGGCGATCGTTAAGAGGCTGCCAACTTCCATCTCTGTTAAGGTGAACAGATCGGTACC 1020
QY 1171 CTCACCGAGACTTCGAGCTGTGACATGCTCACCGGAGGCTACACCTTCCATGATG 1230
DB 1021 CTCACCGAGACTTCGAGCTGTGACATGCTCACCGGAGGCTACACCTTCCATGATG 1080
QY 1231 TCCCAACCGTTCCGGTGAGACCGAGGACACCACTTGTGCTGACCTCGAGTTGCATCAAC 1290
DB 1081 TCCCAACCGTTCCGGTGAGACCGAGGACACCACTTGTGCTGACCTCGAGTTGCATCAAC 1140
QY 1291 TGTGGCCAGATCAAGACTGGTGTCTCCAGCACGTTCCGACCGTGTGCGAAAGTACAACAG 1350
DB 1141 TGTGGCCAGATCAAGACTGGTGTCTCCAGCACGTTCCGACCGTGTGCGAAAGTACAACAG 1200
QY 1351 CTTCTCCGATCGAGCAGCTGTTGGCGACGCGCGCTTACGAGGTCGAGCGCATTC 1410
DB 1201 CTTCTCCGATCGAGCAGCTGTTGGCGACGCGCGCTTACGAGGTCGAGCGCATTC 1260
QY 1411 CCAGCTTTCAAGGGC 1425
DB 1261 CCAGCTTTCAAGGGC 1275
RESULT 11
AX1211169
LOCUS 1275 bp DNA linear PAT 11-MAY-2001
DEFINITION Sequence 1085 from Patent EP1108790.
ACCESSION AX1211169
VERSION AX1211169.1 GI:14037884
KEYWORDS
SOURCE Corynebacterium glutamicum
ORGANISM Corynebacterium glutamicum
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Corynebacteriaceae; Corynebacterium.
REFERENCE
1
AUTHORS Nakagawa, S., Mizoguchi, H., Ando, S., Hayashi, M., Ochiai, K.,
Yokoi, H., Tateishi, N., Senoh, A., Ikeda, M. and Ozaki, A.
TITLE Novel polynucleotides
JOURNAL Patent: EP 1108790-A 1085 20-JUN-2001;
KYOWA HAKKO KOGYO CO., LTD. (JP)
FEATURES
    source
        1..1275
            /organism="Corynebacterium glutamicum"
            /mol_type="unassigned DNA"
            /db_xref="taxon:1718"
ORIGIN
    Query Match 80.7%; Score 1273.4; DB 6; Length 1275;
    Best Local Similarity 99.9%; Pred. No. 2.1e-231;
    Matches 1274; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 151 GTGGCTGAATCATGACGCTATTGCTCGCGAAATCTCGACTCCCGCGGTAAACCAACC 210
DB 1 GTGGCTGAATCATGACGCTATTGCTCGCGAAATCTCGACTCCCGCGGTAAACCAACC 60

```


QY	211	GTGAGGCGAGGTTTCTCGATGAGCGGTTCCTCCACGGTGTCCGAGGTGTTCATCGGC	270
Db	61	GTGAGGCGAGGTTTCTCGATGAGCGGTTCCTCCACGGTGTCCGAGGTGTTCATCGGC	120
QY	271	GCATCCACGGCGTCCACGAGGCTCATGAGTGGTGGAGCGGTGGCATCGCTACCTGGGC	330
Db	121	GCATCCACGGCGTCCACGAGGCTCATGAGTGGTGGAGCGGTGGCATCGCTACCTGGGC	180
QY	331	AAGGGCGTTTGAAGGCGAGTTGAAAACGTCACGAGAAATCGCGACGAGCTCGTGGC	390
Db	181	AAGGGCGTTTGAAGGCGAGTTGAAAACGTCACGAGAAATCGCGACGAGCTCGTGGC	240
QY	391	CTAGAGCTCAGCATCAGCGCTCATCGAGAAATGATCAAGCTTGATGGACCGCC	450
Db	241	CTAGAGCTCAGCATCAGCGCTCATCGAGAAATGATCAAGCTTGATGGACCGCC	300
QY	451	AACAGTCCCGCTGGGTGCAACGCAATCTTGGTGTTCATGGCTGTGCAAGGCT	510
Db	301	AACAGTCCCGCTGGGTGCAACGCAATCTTGGTGTTCATGGCTGTGCAAGGCT	360
QY	511	GCTGCTATTCGCGAGGCTCCCACTGTTCGGCTACATCGGTGACCAAAACGACACGTT	570
Db	361	GCTGCTATTCGCGAGGCTCCCACTGTTCGGCTACATCGGTGACCAAAACGACACGTT	420
QY	571	CTTCAGTTCCAAATGAAATCATCAACGGTGGCGCTCAOGCTCGCTCGGTTGAC	630
Db	421	CTTCAGTTCCAAATGAAATCATCAACGGTGGCGCTCAOGCTCGCTCGGTTGAC	480
QY	631	GTTGAGAAATCATGATCGCTCCAAATCGGTGACAGACCTTCTTGAGGCTCTCCGCAAC	690
Db	481	GTTGAGAAATCATGATCGCTCCAAATCGGTGACAGACCTTCTTGAGGCTCTCCGCAAC	540
QY	691	GGCGCGAGGTCTACACGCACTCAAGTCCGTATCAAGGAAAGGCGCTGTCCACCGGA	750
Db	541	GGCGCGAGGTCTACACGCACTCAAGTCCGTATCAAGGAAAGGCGCTGTCCACCGGA	600
QY	751	CTTCGCAATGAGGCGGCTTCGCTCTTCCGTGGCTCCACCGGTGAGGCTTTGACCTT	810
Db	601	CTTCGCAATGAGGCGGCTTCGCTCTTCCGTGGCTCCACCGGTGAGGCTTTGACCTT	660
QY	811	ATCGTTGGGCAATCGAAGGCTGGCTTACCCGAGGCAAGCATCGCTTGTCTG	870
Db	661	ATCGTTGGGCAATCGAAGGCTGGCTTACCCGAGGCAAGCATCGCTTGTCTG	720
QY	871	GACGTGCTTCTCTGAGTCTTCAAGGCGGCACTTACCACTTCGAGGTGGCGACAC	930
Db	721	GACGTGCTTCTCTGAGTCTTCAAGGCGGCACTTACCACTTCGAGGTGGCGACAC	780
QY	931	TCCGCACTGAGATGGCAACGTTTACGCTGAGCTCGTTGACGCTACCCCAATCGTCTCC	990
Db	781	TCCGCACTGAGATGGCAACGTTTACGCTGAGCTCGTTGACGCTACCCCAATCGTCTCC	840
QY	991	ATCGAGGACCACTGCGAGGAAGATGATGGAGGGTTACCAACCTCACCGGACCATC	1050
Db	841	ATCGAGGACCACTGCGAGGAAGATGATGGAGGGTTACCAACCTCACCGGACCATC	900
QY	1051	GGCGACAGGTTTCAGATCGTTGGGACGACCTTCTCGTCAACACCTTGAGGCTGAAG	1110
Db	901	GGCGACAGGTTTCAGATCGTTGGGACGACCTTCTCGTCAACACCTTGAGGCTGAAG	960
QY	1111	GAGGCGATCGCTAAGAGGCTGCCAATCTCCATCTCTGGTTAAGGTGAACAGATCGGTACC	1170
Db	961	GAGGCGATCGCTAAGAGGCTGCCAATCTCCATCTCTGGTTAAGGTGAACAGATCGGTACC	1020
QY	1171	CTACCGGAGACTTTCAGCGCTGTGACATGGCTACCGCGAGGTACACCTCATGATG	1230
Db	1021	CTACCGGAGACTTTCAGCGCTGTGACATGGCTACCGCGAGGTACACCTCATGATG	1080
QY	1231	TCCACCGCTTCGGGTGAGACCGAGACACCACTTGTGCTACCTCGCAGTTGCACTCAAC	1290
Db	1081	TCCACCGCTTCGGGTGAGACCGAGACACCACTTGTGCTACCTCGCAGTTGCACTCAAC	1140
QY	1291	TGTGGCGAGATCAAGACTGGTGTCTCAGCACGTTCCGACCGTGTGCAAGTACAAACAG	1350
Db	1141	TGTGGCGAGATCAAGACTGGTGTCTCAGCACGTTCCGACCGTGTGCAAGTACAAACAG	1200
QY	1351	CTTCTCCGCGATCGAGCAGCTGCTTGGCGACGCGCGTCTACGAGGTGCGAGCGCATTC	1410
Db	1201	CTTCTCCGCGATCGAGCAGCTGCTTGGCGACGCGCGTCTACGAGGTGCGAGCGCATTC	1260
QY	1411	CCACGCTTTCAGGGC	1425
Db	1261	CCACGCTTTCAGGGC	1275
RESULT 12			
AP005217			
LOCUS			
DEFINITION			
AP005217 Corynebacterium efficiens YS-314 DNA, complete genome, section 4/11.			
ACCESSION			
AP005217 BA000035			
VERSION			
AP005217.1 GI:23492722			
KEYWORDS			
SOURCE			
ORGANISM			
Corynebacterium efficiens YS-314			
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;			
Corynebacterineae; Corynebacteriaceae; Corynebacterium.			
REFERENCE			
AUTHORS			
Nishio, Y., Nakamura, Y., Kawarabayasi, Y., Usuda, Y., Kimura, E., Sugimoto, S., Matsui, K., Yamagishi, A., Kikuchi, H., Ikeo, K. and Gojobori, T.			
TITLE			
Comparative complete genome sequence analysis of the amino acid replacements responsible for the thermostability of Corynebacterium efficiens			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
FEATURES			
source			
1. 300750			
/organism="Corynebacterium efficiens YS-314"			
/mol_type="genomic DNA"			
/strain="YS-314"			
/db_xref="taxon:196164"			
252..620			
/notes="CE0885"			
/codon_start=1			
/transl_table=11			
/product="hypothetical protein"			
/protein_id="BAG17695.1"			
/db_xref="GI:23492723"			
/translation="MDQAPSGCLVDCVHGVIYFDVTPVMLLPVVGTSNKRKIVLDLFP			
HYFLNKPDRILFTPLMSGCGCLFGWIGLFTPHWTHVYVCWFCEPDPGSGRTLAAC			
LTHASRTMKPCLQGLVANG"			
529..1936			
gene			


```

/ gene="rRNA_16S"
529. .1936
/ gene="rRNA_16S"
/ product="16S ribosomal RNA"
/ note="CE0886"
complement (2042. .2215)
/ codon_start=1
/ transl_table=11
/ product="hypothetical protein"
/ protein_id="BAC17696.1"
/ db_xref="GI:23492724"
/ translation="MSSQTSPNSMPTTTHHLLGDVFPADPWVWVHPIITTCWQPHGSGNT
NPPHSHSPWVY"
2399. .5481
/ gene="rRNA_23S"
2399. .5481
/ gene="rRNA_23S"
/ product="23S ribosomal RNA"
/ note="CE0890"
5606. .5725
/ gene="rRNA_5S"
5606. .5725
/ gene="rRNA_5S"
/ product="5S ribosomal RNA"
/ note="CE0890"
5925. .7088
/ note="CE0889, similar to AX063829-1|CAC25156.1| percent
identity: 79 in 387 aa"
/ codon_start=1
/ transl_table=11
/ product="putative aminotransferase"
/ protein_id="BAC17699.1"
/ db_xref="GI:23492727"
/ translation="MRNVVVERLRPFCETIFATWTPORANEAGALNGCGPPDEDGPA
PMLEASQILGGNQSGRGDGLRPTAVASDRHKPKFLDYDPTETVLTVDGATEAI
SASVLGVEPDEVILFPIFYDAYAALALAGATRAVPLKEEDNTWLDITDVAHAAY
TKKTSMLVNSPNSPTGVSFKRSLEALAGARAYDLVLVSDEKVFHTYDPKDSHTAV
ASLPGMDRTITVSSAKPTFNWTKTGWALAPAPILAEAVIRAKFMSYVYGATFPQPA
VAHAHAHEKPPWVKQKGLNKRDLITDALTQAGLKVDHSHGTFFVVAADTGERDGAEP
CFDLIDRVGVAAPVQAPVDPKEQWSKVRPAFCQKQEDTLREAAQRLRAAGSL"
complement (7180. .7821)
/ note="CE0890"
/ codon_start=1
/ transl_table=11
/ product="hypothetical protein"
/ protein_id="BAC17700.1"
/ db_xref="GI:23492728"
/ translation="MTHSFVDDEYARKNNEFLRDTKRLQISAPFFGLVLLAIGGVL
YLLADGAGWMLVLMGIMAAIGFVMIPLVPRQVSAESLYNTYPLVPAVFAEINPR
DVLLALVNTVNTLPPRWGLAVRTINRLGVHDKLGERVPSVAVSGRTIRDQHW
DEISPMPTTGWTDADVVRAAKTIPHDNLNKLKNRNLDEVKNTWMLLVL"
complement (7822. .9495)
/ note="CE0891, similar to AL583924-93|CAC31112.1| percent
identity: 75 in 547 aa"
/ codon_start=1
/ transl_table=11
/ product="putative DNA helicase"
/ protein_id="BAC17701.1"
/ db_xref="GI:23492729"
/ translation="MAFGDGLPIVOSDKTVLLETGHDLAGARAALAPFAELERAPRH
VHTYRTPLALNARAAGHDAEQVDVLYERSFPVQPLLADVATMSYGRVMHK
HPAHLIILSESPAILVEISHRKIKMQLQDLPETIAVHPSERGLKQGLKIGMP
ADBLAGYVDGAHPITALSTVEDSLRDYQYAADSWEGSGVVVLPCCGAKTWGA
ASMAQAQNTLLILVNTVAGQWMDLRLRFTLDELGEYSGERKEIRPVVTIATGVQ
VTRTKGEYKALELPDSRDWGLIYDEVHLLPAPVFRMTSDLSQRRRIGLTAIVRED
GREGVDLSLIGPKRYDAPWKOELAQGFATADCEVETRTMTDAPWYATATADRYR
LAATAHTRVVRVRIEIQOQPTLIIGALDLEELGAEDFADPVGDKTSNKKRGEI
FDRFSGSLNVLVSVKANFSDILPEAAVAIQVSGFSQSEARQLRGLLRPKADGG
BAHFVSVSVRDTLTDEYAHRFLAEGQYAYRIIDADDLLFPITEKET"
complement (9548. .11851)
/ note="CE0892, similar to AE006976-11|AK45126.1| percent
identity: 36 in 741 aa"

```



```

/misc_feature
/locus_tag="DIP0718"
/notes="Doubtful CDS. No significant database matches"
/codon_start=1
/transl_table=11
/product="Putative membrane protein"
/db_xref="GI:38199584"
/translation="MKYRSHNIIVCSLIPQPLKARHIPAKIPSPVRIVELP"
complement(141..206)
/locus_tag="DIP0718"
/notes="Signal peptide predicted for DIP0718 by SignalP 2.0
HMM (signal peptide probability 0.699) with cleavage site
probability 0.667 between residues 22 and 23;
signal-peptide site"
257..484
/locus_tag="DIP0719"
257..484
/locus_tag="DIP0719"
/locus_tag="DIP0719"
/notes="Similar to Mycobacterium tuberculosis CDC1551
hypothetical 10.1 kDa protein MT3304 TR:AAK47646
(EMBL:AE007142) (97 aa) fasta scores: E(): 4.5e-11, 51.38%
id in 72 aa"
/codon_start=1
/transl_table=11
/product="Conserved hypothetical protein"
/protein_id="CAE49241.1"
/db_xref="GI:38199585"
/translation="WDIKIGFSDSPRELVSRSREHNEIVRIHNRALRDAEGLVDLD
DGNRYLVNRNARIAYVEVGSTAPRTVGFAGA"
487..1365
/locus_tag="DIP0720"
487..1365
/locus_tag="DIP0720"
/notes="Similar to Mycobacterium tuberculosis hypothetical
31.0 kDa protein RV3207c or MTCY07D11.19 TR:005859
(EMBL:Z95120) (285 aa) fasta scores: E(): 1.7e-44, 46.6%
id in 287 aa"
/codon_start=1
/transl_table=11
/product="Putative membrane protein"
/protein_id="CAE49242.1"
/db_xref="GI:38199586"
/translation="MTDPHAEFFVRFARDYGRAYAIPLVAVITVWLIDVFRTPAE
TTTTTGGAPTATATACKEGDPARONRPIAIITELPSGPEFTQKRGVTRVGN
GAHAGKHDKVPTLYEVENGINTAYAGGDDAFAMVDATLNRKSWTHDKGPGHEV
DAGAKDPRLQLSDVDTGHLGCGNIAMETSCFIGNRVVINESRWVRGAKFPQG
DIGARKYQLINHEVHGIGIFANHEPCGKNGELAPIMMOQTLSNSLSLFAIDANETYN
DDGAVCSANPWPYPFA"
535..603
/locus_tag="DIP0720"
/notes="1 probable transmembrane helix predicted for
DIP0720 by TMHMM2.0"
1165..1194
/locus_tag="DIP0720"
/notes="ScanRegExp hit to PS00142, Neutral zinc
metallopeptidases, zinc-binding region signature."
1389..2246
/locus_tag="DIP0721"
1389..2246
/locus_tag="DIP0721"
/notes="Similar to Mycobacterium leprae hypothetical
protein ML0818 TR:Q9CCG7 (EMBL:AL583919) (297 aa) fasta
scores: E(): 1.2e-34, 39.47% id in 266 aa"
/codon_start=1
/transl_table=11
/product="Conserved hypothetical protein"
/protein_id="CAE49243.1"
/db_xref="GI:38199587"
/translation="MSLQQLQVHRDAFHAPTPPRLQGVANDYGNRVGNVAFSQVIH
PRNSASRVKDLQPGQFVRVPLKSTDGRIYAGWRNSFEETGLARRVDETVAAL
RLDITALADIDVDFSPEDSHLSADHSAWPEPMVALGVTSDAARVTPALQAL
AQRLPKVILLRDLIDAPIQVTHADMFAITTYSGNMPVVTOLVGVARPYGTAACVI
VDALLAGAVEGIIIDRFSHIQHRDQLLALVYRIALVHALHPESISNTGNLEWVSQT

```

```

gene
2250..5432
/locus_tag="DIP0722"
CDS
2250..5432
/locus_tag="DIP0722"
/notes="Similar to Streptomyces coelicolor putative
ATP-dependent DNA helicase 2SC3B6.07 TR:Q9FCK5
(EMBL:AL390968) (1159 aa) fasta scores: E(): 6.2e-29,
32.21% id in 1161 aa"
/codon_start=1
/transl_table=11
/product="Putative ATP-dependent DNA helicase"
/protein_id="CAE49244.1"
/db_xref="GI:38199588"
/translation="MDAFSTPDQSPPTSAQVIGEDLTAGIPLPQOSFVYQAPRVEIV
YNDHSEPIREWDNQLPQORGTWRVTGAGAGVSTLLMDTVAERTRGVSPESIIVA
ASRESAARLRAGIAHRIADGYSASPASLVRSVHSLAFALRSLSGDEQLRLITGABODA
VIRELQGHVDNPNLSAAMPAAQALGLIGAREVDRDFLLRSRGELPEDELEIGA
RYGRPMWASAGKFMREYQQVMNLGTHSNASELVSKLCLFDIPMGWHTIIDDAQH
LDPSAQLQALMRYTDFVTAGDPPQSFHFRGASPEFLHAPDHRTLYSSPROP
TVEAKVIGSTGQSEFVADLLRRSHLLRGIRWDMVIVRSTADIPSLRRALLSAGVP
VOEDPSDILSEQRIVSALILAVRAIQELSQAQLEELALGPICGADTTLRLFRGL
RKAEMHAGNRRAIEMRELIDPKESESTQLREQVEAVLTDRELAVDKIRAVLRG
AOPGSEVEILWEIWDASLSLSLQTVSLRGVGRGAORDRLDGLVADAGDGVERR
PTASITSEVRHIAEQELPTGVRRLATADAVRIVTAHSGSQQQWHTIVAGVQGTW
PSLOETGTLFGQEDLIDIGSIEPNTYISRAEKLKEKRLFHVTRGTGRVVTI
VESPEDTAAPSPFLOQIAGESKVGQPTTDPETASVQSEIDTETVHLDOYVLLSVPS
IVAEIRRELNPDSFORREQAARQLARLALHQPVGARPEQWVGGSSETSLDISK
VPSLIENALICPLARLERLVEBENTPIHMLKGTLAHAFABEAVGGRVDPNBAQLVT
QAFEALLDAPAMSLPHHMSOWSTMQLRAHMDVSHAQRLVGVVPPVNVVAVGVEL
RGRIDRLERNDAGEFHVDFDKTGQAVTKDBANENKQLLAYQLALHRLGKLMORNGEPA
INTTAEOPLGLTVDOAVLVYPATDKTVTTRQAPKDAEELKFSATLPALLSLSGPQ
LVARINFCDOCKIKTMCPAQPEGKWPEC"
2250..2558
/locus_tag="DIP0722"
/notes="ScanRegExp hit to PS00430, TonB-dependent receptor
proteins signature 1."
3045..3200
/locus_tag="DIP0722"
/notes="HMMpfam hit to PF00580, UvrD/REP helicase"
3285..3431
/locus_tag="DIP0722"
/notes="HMMpfam hit to PF00580, UvrD/REP helicase"
4167..4199
/locus_tag="DIP0722"
/notes="ScanRegExp hit to PS00136, Serine proteases,
subtilase family, aspartic acid active site."
5426..8656
/locus_tag="DIP0723"
5426..8656
/locus_tag="DIP0723"
/notes="Similar to Mycobacterium tuberculosis CDC1551
helicase, UvrD/Rep family MT3295 TR:AAK47638
(EMBL:AE007142) (1101 aa) fasta scores: E(): 2.9e-25,
43.03% id in 1120 aa"
/codon_start=1
/transl_table=11
/product="Putative helicase"
/protein_id="CAE49245.1"
/db_xref="GI:38199589"
/translation="MLSPQELSCALGOKFPPTPOADVSSPLAPTLVAGAGAKTE
TMAARVNLVAGLVDPDRVLGLTFRKAAQOLSKRDLREQLAGINLDRDPTGA
LATKLEAIPVTYSDYAGRLISEYGLLLPFESSRLISQITELFQIAHSIVSAHTGA
LNTSNAENTVSTLISLVSEMDNEMVSPDIEEBSAAPLAMIETVSKRAPSKEVY
KMRDYOIRNELLPLVQLKTHLADNHLMTFGEQMSLAARAAENPQVGSQRNRYQI
IMLDEYODTGHQARVLLKSLFAGTAVTAVGPMQSIYGRGATANLIERFLTDFGNSG
SPATKELTVSFRNPPVELDIANFVSRELICVPEDPREPQPLEGPAALNGIVRLGF
PPSNDERSYVADHLAAQYEGHDGHPPTAAVIRKRXHSAIALELQORGVPVEIIG
LGLGIPVADHLAIAITLLVRPYDTQAMRILAGSVGLGADMLALMALDRYNLSGR
DRRTATLRLDRPLERLKOIADTTPSQDSIVTGLAEAVDLDERDSDGPRYSAGKS
ERLTATLRLRNLSTNSLSLPLDFADIERVFGITRTEVLDQRPDSGATGTAKHLD
FAEVQDFSRIPGANLSLLDYLSLAESENGLEPGEVQVTAQVQILTUVHAKGLEW
QHVAVLHADANTYVAKASTWLTNANASVPSALRGDKGDEDLVGAPVEIDTPTAAEL

```

AKAGAHIAIDPKQVAENARLFVVAITRAEQVLVITASDPKRPVLPEYELTMLR
NDPDSVEWHERGEADYPPAPQEAFFPNYIVGAEDVFAAMQKQDPLISDDDLF
ERMEKVSALIEEHEQLSAPVAVNIGVELTATDINLAKNPENFAQRERPPVFPKP
SYAKRGTAFTHEMLNRFCAALDELTELPGIBELDDLDRLKEAPFLDSEWADRTPE
HVEHPFVSIGRHIVRGMADVFNKSDGRWIVVDKTPQPTDNHKEKESVAMQLAVR
LAWARLQIDVDRVDAVPHYVGRNRYRPOQLPDGBELALKLDPVS"
5474..6931
/locus tag="DIP0723"
/note="HMPfam hit to PF00580, UvrD/REP helicase"
5531..5554
/locus tag="DIP0723"
/note="ScanRegExp hit to PS00017, ATP/GTP-binding site
motif A (P-loop)."
8701..9789
/locus tag="DIP0724"
8701..9789

misc_feature
misc_feature
gene
CDS

Query Match 58.0%; Score 915.8; DB 1; Length 347625;
Best Local Similarity 81.4%; Pred. No. 1e-163;
Matches 1061; Conservative 0; Mismatches 242; Indels 0; Gaps 0;

147 CACAGTGCTGAATCATGACGATTCGCTCGCGAAATTCCTGACTCCCGCGGTAAACCC 206
Db 201513 CACAGTGCTGACATTAAGCAGATTAATTCCTGCGAAATTCCTGACTCCCGCGGTAAACCC 201572

207 AACCGTCGAGGACAGAGGTTTCTCGATGACGCTTCCACCGGTGTCGAGGTGTTCCATC 266
Db 201573 AACCGTTGAAGCGAAAGTTTCTTGGATGACGGATCCACCGGTGTCGCGGTCTCTTC 201632

267 CGCGGCATCCACCGGCTGCCAGAGCTCATGAGTTCGGTGACGGTGCGATCGCTACCT 326
Db 201633 CGGTGCTGCCACCGGTGTTTCAAGAGCTTCAAGAGCTTCCGACGGTGCGAGCGCTACCT 201692

327 GGGCAAGGCGTTTGAAGGAGTTTGAAGAGTGAAGAGAGTGAAGAGTGAAGAGTGAAGAGT 386
Db 201693 GGGCAAGGCGTTTCAATGAGTTTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGT 201752

387 TGGCTAGAGGTGACGATGACGCTCATCATCAAGAGCAATGATCAAGTTGATGCGAC 446
Db 201753 TGGTCAGAGCGGACGATGACGCTCATCATCAAGAGCAATGATGATGCTCGACGCGAC 201812

447 CGCCAAAGTCCCGGCTGGGTGCAACGCAATCTTGGTGTTCCTGATGCTGTGCTGTGCAA 506
Db 201813 TGAGAACAGTCTGCTCTCGCGCAAGAGCTATCTTGGGTGTATCTCATCGCGGTAGCTAA 201872

507 GGCTGCTGATTCGCGAGCGCTCCACTGTTCCGCTACATCGGTGACCAAGAGCA 566
Db 201873 GGCTGCCGAGATGCTGCTGCGCTACCTTTGACCGCTACATCGCGCGCTTACGCTCA 201932

567 GGTCTTCCAGTTCCAAATGATGAACATCATCAAGGTGGCGCTCACGCTGACTCGGTGT 626
Db 201933 GGTCTTCCAGTTCTTATGATGAACATTTGTTAAAGTGGCGCACAGCTGACTCGGCT 201992

627 TGACCTTCCAGGATTCATGATGCTTCCAAATCGGTGACAGACCTTCTCTGAGGCTCTCG 686
Db 201993 TGATGTTCCAGGATTCATGATGCTTCTTATCGGTGCGAGTCTTCTCTGAGGCTCTCG 202052

687 CAACGGCGGAGGTCTACCGACGCTCAAGTCCGCTCATCAAGGAAAGGCGCTGTCCAC 746
Db 202053 CATGGGTGACAGGCTTACCTCTCTTGAAGTCCGCTGATTAAGTCAAGGAGCTTTCAC 202112

747 CGGACTTGGCGATGAGGCGGCTTCGCTCTTCCGTCGCTCCACCGGTGAGGCTCTTGA 806
Db 202113 CGGCTTCGGGACGAGGTGGTTTTCGACCTTCTGTTAGTCCACCAAGGCGAGCTCTGA 202172

807 CTTATCGTTAAGGCAATCGAGAAGCTGGCTTCAACCGGCAAGGACATGCTCTTGC 866
Db 202173 CTTATCGTTAAGGCAATCGAGAAGCTGGCTTCAACCGGCAAGGACATGCTCTTGC 202232

867 TCTGACGCTTCTCTCTGAGTCTTCAAGGACGCGACTTACCACTTTCGAAGGTGGCA 926
Db 202233 ACTCGAGTGTCTTCTCTGAGTCTTCAAGGATGCGAAGTACCACTTTCGAAGGTGGCA 202292

QY 927 GCACCTCCGAGCTGAGATGGCAACGTTTACGCTGAGCTCGTTGAGCGCTACCCCAATCCT 986
Db 202293 GCACACCGCTGAGGAGATGGCAAGGCTTACGAGGAGCTCATCGCTGAGTACCCCAATCT 202352

QY 987 CTCCTATCGAGGACCCCACTCAGAGAGATGACTGGGAGGGTTACACCAACCTCACCGCAAC 1046
Db 202353 TTCCTATCGAGGACCCCACTCAGAGAGAGAGCTGGGAGGGCTACACCGCTTACCGCGCG 202412

QY 1047 CATCGCGCAAAAGTTTCAGATCGTTGGCGAGAGCTTCTTCGTCACCAACCTTGGAGCGCT 1106
Db 202413 AATCGGTGACAAGTTTCAGATCGTTCGCGAGAGCTTCTTCGTCACCAACCTTGGAGCGCT 202472

QY 1107 GAAGGAGGATCGCTTAAGAGCTGCCAATCCATCTCTGTTAAAGTGAAGTGAAGTGAAGTGAAG 1166
Db 202473 CAAGGAAGGATCGAGAGAGGCTGCCAATGCTTCTGCTGGAAGTGAAGTGAAGTGAAGTGAAG 202532

QY 1167 TACCTCTCACGAGACCTTCGAGCTGTCGACATGCTTCACCGCGCAGGCTTACACCTCCAT 1226
Db 202533 TACCTCTCACGAGACCTTCGAGCTGTCGACATGCTTCGCGACCGCAACCGCTACCGACCAT 202592

QY 1227 GATGTCCCAACCGTTCCGTTGAGACCGAGGACACCAATTCGTCGACTCGAGTTCGACT 1286
Db 202593 GATGTCCCAACCGTTCCGTTGAGACCGAGGACACCAATTCGTCGACTCGAGTTCGACT 202652

QY 1287 CAATGTGGCAGATCAAGTCTGCTCCAGCAGCTTCGAGCGCTTCGAGCGGTGTCGAAAGTACAA 1346
Db 202653 GGGCTGTGGCAGATCAAGCAGCTTCGAGCAGCTTCGAGCGGTGTCGAGCGGTGTCGAAAGTACAA 202712

QY 1347 CGAGCTTCTCCGATCGAGCAGCTGCTTGGCGACCGCGCTCTAGCGAGGTGTCGAGCGC 1406
Db 202713 TGAGCTTCTCCGATCGAGCAGCTGCTTGGCGACCGCGCTCTAGCGAGGTGTCGAGCGC 202772

QY 1407 ATTCCCAACCGTTTCAGGGCTTAAATAAAGCGCTTTTCGAGCGC 1449
Db 202773 ATTCCCAACCGTTTCAGGGCTTAAATAAAGCGCTTTTCGAGCGC 202815

RESULT 14
SCO939115/c
LOCUS
DEFINITION Streptomyces coelicolor A3(2) complete genome; segment 12/29.
ACCESSION AL939115 AL035654 AL035707 AL049573 AL132674 AL354048
AL35832 AL359989 AL442120 AL596138 AL645882
VERSION AL939115.1 GI:24426505
KEYWORDS
SOURCE Streptomyces coelicolor A3(2)
ORGANISM Streptomyces coelicolor A3(2)
REFERENCE 1
AUTHORS Bentley, S.D., Chater, K.F., Cerdano-Tarraga, A.M., Challis, G.L.,
Thomson, N.R., James, K.D., Harris, D.E., Quail, M.A., Kieser, H.,
Harper, D., Bateman, A., Brown, S., Chandra, G., Chen, C.W., Collins, M.,
Croft, A., Fraser, A., Goble, A., Hidalgo, J., Hornby, T., Howarth, S.,
Huang, C.H., Kiser, T., Larke, J., Murphy, L., Oliver, K., O'Neill, S.,
Rabinowitz, B., Rajandream, M.A., Rutherford, K., Rutter, S.,
Seeger, K., Saunders, D., Sharp, S., Squares, R., Squares, S.,
Taylor, K., Warren, T., Wietzorrek, A., Woodward, J., Barrall, B.G.,
Parkhill, J. and Hopwood, D.A.
Complete genome sequence of the model actinomycete Streptomyces
coelicolor A3(2)
JOURNAL Nature 417 (6885), 141-147 (2002)
MEDLINE 21996410
PUBMED 12000953
REFERENCE 2 (bases 1 to 276800)
AUTHORS Bentley, S.D.
TITLE Direct Submission
JOURNAL Submitted (09-MAY-2002) Submitted on behalf of the Streptomyces
sequencing team, Sanger Institute, Wellcome Trust Genome Campus,
Hinxton, Cambridge CB10 1SA E-mail: sds@sanger.ac.uk
On or before Oct 30, 2002 this sequence version replaced
gi:4500374, gi:4490616, gi:4490978, gi:6138834, gi:7672242,
gi:7799503, gi:6894754, gi:10241774.


```
gene
/notes="Degenerately, directly repeated at 12216..13297"
complement(4025..4567)
/genes="SCO3059"
CDS
/notes="synonyms: purE, SCBAC19G2.14c"
complement(4025..4567)
/genes="SCO3059"
/BC number="4.1.1.21"
/notes="SCBAC19G2.14c, phosphoribosylaminoimidazole
carboxylase catalytic subunit PurE, len: 180aa: strongly
similar to many eg. SW:Q44679 (PurE, CORAM)
phosphoribosylaminoimidazole carboxylase catalytic subunit
PurE from Corynebacterium ammoniagenes (177 aa) fasta
scores: opt: 666, z-score: 735.8, 65.541% identity
(65.541% ungapped) in 148 aa overlap and TR:080937
(EMBL:AC004684) putative phosphoribosylaminoimidazole
carboxylase from Arabidopsis thaliana (645 aa) fasta
scores: opt: 644, z-score: 703.2, 60.870% identity
(60.870% ungapped) in 161 aa overlap. Contains Pfam match
to entry PF00731 AIRC, AIR carboxylase."
/codon_start=1
Query Match 46.4%; Score 732.4; DB 1; Length 276800;
Best Local Similarity 72.5%; Pred. No. 6.4e-129;
Matches 962; Conservative 0; Mismatches 361; Indels 3; Gaps 1;
QY 106 CGTGAATAATCTAGTTAGTCCCAAGTTCGCATAGGAGGCCACAGTGGCTGAATCATG 165
DB 43666 CGTGAAGCGTCACAGTGAAGCGTCACATTAAGAAGGAGATGCTGTGCGCTCANTCGAC 43607
QY 166 CACGATATTCGCTCCGGAATTTCTGACTCCCGCGGTAAACCAACCGTCGAGGAGAGGTT 225
DB 43606 GTGCTGTAGCCCGGGAATCTCTGACTCCCGAGGCAACCCACGGTCGAGGTCGAGTTC 43547
QY 226 TTCTCGATAGCGTTTCCACAGTGTGCGAGTGTTCATCCCGCGCATCCACCGGGTTC 285
DB 43546 GGCCTCGACAGCGCAGCAGGGTGTGTCGCGCCGCTTCCGTCGCGGCGCTCCACCGGCGCC 43487
QY 286 CACGAGCTATGAGCTGCGTGAAGTGGC---GATCGTACCTGGCGCAAGGGCGTTTTC 342
DB 43486 TTCGAGGCCATCGAGCTGCGTGAAGCGGCAACCGAGCGCTTACCTCGGCAAGGGCGTTCGAG 43427
QY 343 AAGCAGTTGAACAGCTCAACGAAGAAATCGGCAACGAGCTCGCTGGCCCTAGAGGCTGAC 402
DB 43426 AAGCCGTGTGCGCGTCTATCGACAGATCGGCCCGGAGCTGCTGCTGCGTACGAGCGCCACC 43367
QY 403 GATCAGCGCTTCATCGACGAAGCAATGATCAAGTGTGATGGCACCGCCAAAGTCCCGC 462
DB 43366 GAGCAGCGCTGATCGACAGGCCATGTTTCGACCTGGACGCGCACCGCAACAGGGCTCG 43307
QY 463 CTGGGTGCAACGCAATCTTGGTGTTCATGCTGTTGCAAGGCTGCTGCTGATTCC 522
DB 43306 CTCGGCGCAACGCCATCTCGGGTGTCTCCCTCGCGGTGCGCCACGCGCGCTCCGAGGCC 43247
QY 523 CGAGCGCTCCCACTGTTCCGCTATCATCGGTGGAACCAACGACACAGTTCCTTCCAGTTCCA 582
DB 43246 AGCGACCTGCGCTCTTCGCTACTCTGGCGGCGCGCAAGCGACCTGCTGCGGTGCGG 43187
QY 583 ATGATGAACATCATCAACGGTGGCGCTCACGCTGATCTCCGTTGATGAGTTCAGGAATTC 642
DB 43186 ATGATGAACATCTGAACGGGCGCTCGCACGCGGACTTCAACGCTGGAATCCAGGAGTTC 43127
QY 643 ATGATCGTCCAACTCGGTGCAAGACCTTCTGAGGCTCTCCGCAACGCGCGCGAGGTC 702
DB 43126 ATGATCGCCCCGATCGGCGGAGTCTCTTCGAGGCGTCTGCTGCGGCGCGCGAGTTC 43067
QY 703 TACCACGCACTGAAGTCCGTCATCAAGGAAAGGCGCTGTCCACCGGACTTTGGCGATGAG 762
DB 43066 TACCACACCTCAAGAGTCTCTGAAGAACAGGGCTTGGCCACCGGCTCGGCGACGAG 43007
QY 763 GCGCGCTTCGCTTCCTCGTGGCTCCACCGTGAAGGCTCTTGACCTTATGTTAAGCA 822
DB 43006 GCGCGCTTCGCGCCGAAACCTGGGCTCCAAACCGGAGGCGCTCGAGCTCATCTCGAGGCG 42947
```

```
QY 823 ATCAGAGAAGGTGCTTCAACCCAGGCAAGACATCGCTTGTGCTTGGAGCTGTGCTTCC 882
DB 42946 ATCAAGGAGGCGCGCTACACCCCGGCGAGCAGATCGCCCTCGCGCTCGACGTGCGCGC 42887
QY 883 TCTGAGTTCCTCAAGGAGCGGCACTTCACTTCCAGGTGCGCAGCACTCCGCGAGCTGAG 942
DB 42886 TCCGAGTTCCTCAAGGAGCGGCTCTTACGCTTTCGAGGGCAAGAACCGCTCCGCGCGCAG 42827
QY 943 ATGCAAAAGCTTTACGCTGAGCTGTTGACGCGTACCCAACTCGTCTTCCATCGAGGACCA 1002
DB 42826 ATGACCGAGTACTACGCGAGCTGCTGAGGGGTACCCGCTGCTTCCATCGAGGACCG 42767
QY 1003 CTGCAAGGAGATGATCTGGAGGGTTACCAAACTCACCAGCAACATCGCGGACAGGTT 1062
DB 42766 CTGTTTCGAGGAGCAGCTGGGAGCGGTGGAAACCACTCACCGCAAGCTCGGCGCAAGGTG 42707
QY 1063 CAGATCGTTGGGAGCAGCTTCTTGTGTCACCAACCTGAGCGCTGAGAGGAGGCAATCGCT 1122
DB 42706 CAGCTGGTGGGAGCAGCTGTTGCTTCCAAACCGGAGCGCTTGGCCGCGGATCGAG 42647
QY 1123 AAGAAAGGCTGCCAACTCCATCTGTTAAGGTGAACCAAGATCGGTACCTCACCGAGACC 1182
DB 42646 GGAAGCTCGGCAACGCGCTGCTGCTCAAGGTCAACCAAGATCGGTTCGCTGACCGAGACT 42587
QY 1183 TTGCAAGCTGTGCAATGGCTCACCGGAGGCTACACCTCACTGATGATGATGATGATGATG 1242
DB 42586 CTGCAAGCGCTGAGCTGGCCAGCGCAACGCGCTTCAAGTGCATGATGATGATGATGATG 42527
QY 1243 GGTGAGACGAGGACCAACCACTTCTGCTGACCAACCTGAGCGCTGAGAGTTCAGTTCGCGCAGATC 1302
DB 42526 GCGGAGACGAGGAGCTGACCACTCCGACCTGGCGCTGCGCAGCTGCGCAGATC 42467
QY 1303 AAGACTGGTGTCTCCAGCAGCTTCCGACCGTGTGCAAAAGTACAAACAGGTTTCCGCGATC 1362
DB 42466 AAGACCGGCGCGCGCGCTCCGAGCGGTGCGCAAGTACAAACAGCTGCTGCGCATC 42407
QY 1363 GAGCAGCTGTTGGGACGCGCGGCTTACGAGGTTCGAGGCGATTCGAGGCGATTCGCGCTTTCAG 1422
DB 42406 GAGGAGATCTCGACGACCGCGGTGTACGCGCGCGCTTCCCGCGCGCTTCCCGCTTCAAG 42347
QY 1423 GCGTAA 1428
DB 42346 GCGTGA 42341
RESULT 15
AB016822_17/c
WPCOMMENT
Sequence split into 26 fragments LOCUS AB016822 Accession AB016822
Fragment Name Begin End
AB016822_00 1 110000
AB016822_01 100001 210000
AB016822_02 200001 310000
AB016822_03 300001 410000
AB016822_04 400001 510000
AB016822_05 500001 610000
AB016822_06 600001 710000
AB016822_07 700001 810000
AB016822_08 800001 910000
AB016822_09 900001 1010000
AB016822_10 1000001 1110000
AB016822_11 1100001 1210000
AB016822_12 1200001 1310000
AB016822_13 1300001 1410000
AB016822_14 1400001 1510000
AB016822_15 1500001 1610000
AB016822_16 1600001 1710000
AB016822_17 1700001 1810000
AB016822_18 1800001 1910000
AB016822_19 1900001 2010000
AB016822_20 2000001 2110000
AB016822_21 2100001 2210000
AB016822_22 2200001 2310000
```

QY	1097	GTGAGCGCTGAAGGAGGCGATCGCTAAGAAAGGTCGCAACTCATCTCTGGTTAAGGTGA	1156
DB	82149	CGAAGCGCCTTGGCGCGCAGGGCATCCAGCAGAAAGCCGCGAACAGCATCTCTGTTCAAGGTGA	82090
QY	1157	ACCAGATCGGTACCTCTCACCGAGACCTTTCGACGCTGTCGACATCGCTCACCGCGCAGGCT	1216
DB	82089	ACGAGATCGGCACCTTGAACCGAGACGCTTGAACGGGCTCTCGCTCGCCGAGCGCAGCGTA	82030
QY	1217	ACACTCTCATGATGTCCACCGCTTTCGGTTCGAGACCGAGGACACCACTTGTCTGACCTCG	1276
DB	82029	TGACCGCGCTCTCTCTCGCAGCGCTTCGGCGAGACCGAGGACACCACTCGCCGACCTCG	81970
QY	1277	CAGTTGCACCTCAACTGTGGCCAGATCAAGACTGGTGTCTCCAGACGTTTCCGACCGTCTCG	1336
DB	81969	CCGTCCGACCGATTCGCGCCAGATCAAGACCGGCGCCCGGCGCTTCGAGGCGCTCG	81910
QY	1337	CAAAAGTACAAACAGCTTCTCCGCACTCGACGAGCTGTGTCGACGCGCGGCTCTTACGCAG	1396
DB	81909	CGAAGTACACCACTGCTCTCCGCACTCGAGGAGGAGCTTGGGCGAGGCGCTGTCTTACGCCG	81850
QY	1397	GTCCGAGCGCATTTCCCAACGCTTTTCAGGGCTAA	1428
DB	81849	GCCGTTCCGGGTTCCCGCTTACCAGGGCTGA	81818
Search completed: September 28, 2005, 20:45:27			
Job time : 6714 secs			

AE016822_23	2300001	2410000	Continuation (18 of 26) of AE016822 from base 1700001 (AE016822 Leifsonia xyli subsp. xy)
AE016822_24	2400001	2510000	
AE016822_25	2500001	2584158	
	2500001	2584158	
Query Match			
Best Local Similarity 44.4%; Score 700; DB 1; Length 110000;			
Matches 929; Conservative 0; Mismatches 360; Indels 3; Gaps 1;			
QY	140	AGGAGGCCACAGTGGCTGAAATCATGCAGGTATTTCGCTCGGAAATTCCTCGACTCCCGCG	199
DB	83109	AGGAGACTGTTGTGGCTGCANTCGAAGCTGTAGGAGCTCGGAGATCTCTCGACTCGCGG	83050
QY	200	GTAACCCACCGTTCGAGGACAGGTTTCTGATGACGGTTCCACGGTTCGCGAGGTG	259
DB	83049	GAACCCGACCGTTCGAGGTTCGAGTCTCTCGAAGATGCGCAGGTTCAGCCGCGCGCG	82990
QY	260	TTCCATCCGGCGCATCCACCGGCTCCACGAGGCTCATGAGCTTCGCTGACGGTGGC--G	316
DB	82989	TCCCGTCCGGCGGTCCACGGCGGCTTCGAGGCTTACGAGCTTCGCGACGGCGCAAGG	82930
QY	317	ATCGCTACTTGGCAAGGCGGTTTTCGAGGCGAGTTGAAAACGTCACAGAAATTCGGCG	376
DB	82929	GCGCTACTTGGCAAGGTTTCGAGAGGCGCTGAGTGCCTGCTCGACGAGATTCGGC	82870
QY	377	ACGAGCTCGCTGGCTTAGGCTGACGATCAGCGCTCATCGACGAGCAATGATCAAGC	436
DB	82869	CGGCCATTCGAGGCTTCGAGGCGCAGGACCGCGCATCGTCGATGAGGCGATGATCGAAC	82810
QY	437	TTGATGGCACCGCCAAAGTCCCGCTGGGTGCAAAACGCAATCTTGGTTCATCGG	496
DB	82809	TCGACGGCACCGACAAACAAGAACGCGCTGGCGCTAACGCCATGCTCGGTGTGAGCTCG	82750
QY	497	CTGTTGCAAGGCTGCTGTGATTCGCGAGGCTCCACGCTTCGCTACATCGGTGGAC	556
DB	82749	CAGTCGGAAGCGCGCGCGACTCCGCGGACTGCGCGCTGTTCCGCTACCTTGGGCGGC	82690
QY	557	CAAGCGCACAGTTCCTCCAGTTTCCAATGATGAACATCATCAACGGTGGCGCTCACGCTG	616
DB	82689	CGAACCGCATGTCCTGCCGCTCCCGATGATGAACATCATCAACGGCGCGCGCACGGG	82630
QY	617	ACTCCGCTGTGAGCTTCAGGAATTCATGATCGCTCCAATCGGTGAGAGACTTCTCTG	676
DB	82629	ACACCGGCTCGACATCCAGGATTCATGATTCCTCCGATCGGCGCGGAGACTTCTCG	82570
QY	677	AGGCTCTTCGCAACGGCGCGAGGTCTACGACGACTGAAGTCCGTCATCAAGGAAAGG	736
DB	82569	AGGGCTTGCCTGGGGCGTGGAGACCTTACCACTCGCTCAAGGCGCTGCTGAAGTCGAAGG	82510
QY	737	GCTGTCCACCGGACTTGGCGATGAGGCGGCTTCGCTCTTCGTCGGCTCCACCCGCTG	796
DB	82509	GCTTGAACACCGGCTGGGCGACGAGGCGGCTTCGCCCGGAGCTGGAGCACACCGCG	82450
QY	797	AGGCTCTTGACCTTATCGTTAAGGCAATCGAAGGCTGGCTTCAACCCAGGCAAGGACA	856
DB	82449	CTGGCTCGACCTGATCGCGGAGGCCATCGAAGGCGCGCTTCAACGTCGGCTCGAGA	82390
QY	857	TGCTCTTGTCTGAGAGTTGCTTCTCTGAGTTCTTCAAGGACGGCACCTACCACTTCG	916
DB	82389	TCGCGCTCGGCTTCGACGTTGGCTCCACCGAGTTCTTCGAGAACGGCGTCTACCGCTTCG	82330
QY	917	AAGGTGGCGAGCACTCCGAGCTGAGATGCAAAACGTTTACGCTGAGCTCGTTGACGGT	976
DB	82329	AGGGGAGGACCGCACCGCGCGCGAGATGAGCGCTTACTACCGATCTCGGAAACAAC	82270
QY	977	ACCCAAATCGTCTTCATCGAGGACCCACTGCGAGGAGATGACTGGAGGGTTACACCAACC	1036
DB	82269	ACCGCTGTTCTCCTTCAGGACCGCTGCGGAGGACGACTGGGAGGGCTGGGCGCAC	82210
QY	1037	TCACCGCAACCATCGGCGACAAGGTTTCAGATGTTGGCGAGCACTTCTTCGTACCAACC	1096
DB	82209	TGAACCGGAGATCGGCTCGACGCTGCAACTCGTCGGCGATGACCTGTTCTCGTCAACCAACC	82150

Search completed: September 28, 2005, 20:45:27
Job time : 6714 sec

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2005, 23:24:17 ; Search time 93.5 Seconds
(without alignments)
1758.002 Million cell updates/sec

Title: US-10-728-947-4
Perfect score: 2155
Sequence: 1 VABIMHVAREILDRGNPT.....QLLGDAVYAGRSAPPRFQG 425

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:.*
1: Geneseqp1980s:.*
2: Geneseqp1990s:.*
3: Geneseqp2000s:.*
4: Geneseqp2001s:.*
5: Geneseqp2002s:.*
6: Geneseqp2003as:.*
7: Geneseqp2003bs:.*
8: Geneseqp2004s:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2151	99.8	425	4 AAB70882	Aab70882 C. glutam
2	2148	99.7	425	4 AAG90831	Aag90831 C. glutami
3	2148	99.7	425	6 ABP97001	Abp97001 Coryneb
4	2145	99.5	425	4 AAB79278	Aab79278 Coryneb
5	2145	99.5	425	7 ADD13323	Add13323 C. glutam
6	2144	99.5	425	6 ABP97002	Abp97002 Modified
7	1913	88.8	425	6 ABU25984	Abu25984 Protein e
8	1562	72.5	423	8 ADN25687	Adn25687 Bacterial
9	1560.5	72.4	429	6 ABU34209	Abu34209 Protein e
10	1553.5	72.1	429	6 ABU36566	Abu36566 Protein e
11	1553.5	72.1	439	6 ABU34772	Abu34772 Protein e
12	1505.5	69.9	455	4 AAU54498	Aau54498 Propionib
13	1505.5	69.9	455	6 ABM51017	Abm51017 Propionib
14	1502.5	69.7	447	6 ABU35758	Abu35758 Protein e
15	1422	66.0	423	8 ADS22663	Ads22663 Bacterial
16	1412.5	65.5	426	8 ADS27878	Ads27878 Bacterial
17	1408	65.3	429	6 ABP58571	Abp58571 Thermoana
18	1402.5	65.1	423	8 AD21524	Ad21524 Bacterial
19	1402.5	65.1	426	8 ADS28853	Ads28853 Bacterial
20	1402	65.1	429	8 ADS28400	Ads28400 Bacterial
21	1396	64.8	424	8 ADS22786	Ads22786 Bacterial
22	1394	64.7	420	8 ADS25706	Ads25706 Bacterial
23	1394	64.7	423	8 ADS26189	Ads26189 Bacterial
24	1392.5	64.6	431	6 ABU24836	Abu24836 Protein e
25	1377.5	63.9	430	8 ADO59753	Ado59753 B. subtil

ALIGNMENTS

RESULT 1

AAB70882
ID AAB70882 standard; protein; 425 AA.

XX AC AAB70882;

XX 12-JUL-2001 (first entry)

XX C. glutamicum enolase protein.

XX Enolase; fermentation; L-amino acid; L-lysine; coryneform; eno gene;
XX medicine; animal feed supplement.

XX Corynebacterium glutamicum.

XX EPI090998-A1.

XX 11-APR-2001.

XX 29-SEP-2000; 2000EP-00121158.

XX 05-OCT-1999; 99DE-01047791.

XX (DEGS) DEGUSSA-HUELS AG.

XX Moeckel B, Pfeifferle W, Hermann T, Puehler A, Kalinowski J;
XX Bathe B;

XX WPI; 2001-292928/31.

XX N-PSDB; AAF61696.

XX New enolase gene from coryneform bacteria, used to prepare transformants
XX with increased synthesis of amino acids, particularly lysine.

XX Claim 6; Page 14-16; 25pp; German.

XX This invention describes a novel isolated nucleic acid (I) from
XX coryneform bacteria which is used in a method for fermentative production
XX of L-amino acids, especially L-lysine, by fermenting a lysine-producing
XX coryneform in which the eno gene has been amplified, and isolating amino
XX acids that have accumulated in the medium or cells. (I), which encodes an
XX enolase, is used to transform coryneforms for production of L-amino
XX acids, specifically lysine which is used in medicine and particularly as
XX animal feed supplement. It may also be used as probes and primers for
XX isolating related sequences. Overexpression of (I) improves production of
XX amino acids, especially of L-lysine. This sequence represents the enolase
XX protein described in the method of the invention

Ad444869 Bacterial
Abu18243 Protein e
Abu25001 Protein e
Ad433300 Bacterial
Ad442576 Bacterial
Abu48815 Listeria
Abu33030 Protein e
Ad430458 Bacterial
Abu23996 Protein e
Abu19880 Protein e
Adn26444 Bacterial
Adn25109 Bacterial
Adn2035 Bacterial
Abu21358 Protein e
Adn24794 Bacterial
Adn24418 M. methyl
Aau33516 Enterococ
Aau35261 Enterococ
Abu14630 Protein e
Adn17379 Bacterial

26 1377.5 63.9 430 8 ADS44869
27 1373 63.7 431 6 ABU18243
28 1366.5 63.4 448 6 ABU25001
29 1362.5 63.2 426 8 ADS43300
30 1362.5 63.2 427 8 ADS42576
31 1359.5 63.1 430 5 ABB48815
32 1359.5 63.1 430 6 ABU33030
33 1357.5 63.0 430 8 ADS30458
34 1354.5 62.9 431 6 ABU23996
35 1351.5 62.7 427 6 ABU19880
36 1347.5 62.5 417 8 ADN26444
37 1347.5 62.5 424 8 ADN25109
38 1347.5 62.5 426 8 ADN2035
39 1347.5 62.5 427 6 ABU21358
40 1347.5 62.5 427 8 ADN24794
41 1342.5 62.3 428 8 ADR44418
42 1342 62.3 429 4 AAU33516
43 1342 62.3 432 4 AAU35261
44 1342 62.3 432 6 ABU14630
45 1334.5 61.9 426 8 ADN17379

PT mutation point of a gene, measuring expression of a gene, analyzing
PT expression profile or pattern of a gene and identifying homologous gene.
PS Claim 17; SEQ ID NO 4585; 246pp + Sequence Listing; English.
XX
XX The present invention provides a number of nucleotide and protein
CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These
CC are useful for identifying the mutation point of a gene derived from a
CC mutant of coryneform bacterium, measuring expression amount and analyzing
CC the expression profile or expression pattern of a gene derived from
CC Coryneform bacterium, and identifying a homologue of a gene derived from
CC Coryneform bacterium. Coryneform bacteria are useful for producing amino
CC acids, nucleic acids, vitamins, saccharides and organic acids, in the
CC particularly L-lysine. The present sequence is a protein described in the
CC exemplification of the invention. Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from the European Patent Office
XX
SQ Sequence 425 AA;
Query Match 99.7%; Score 2148; DB 4; Length 425;
Best Local Similarity 99.5%; Pred. No. 4.5e-169;
Matches 423; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 VAEIMHVFAREILDSRGNTVEAEVFLDDGSHGVAGVPSGASTGVHEAHLRDGDRYL 60
Db :|||||
Qy 1 MAEIMHVFAREILDSRGNTVEAEVFLDDGSHGVAGVPSGASTGVHEAHLRDGDRYL 60
Db :|||||
Qy 61 KGVLKAVENNEEIGDELADDDORLIDEMIKLDTANKSRLGANAILGVSMVAKA 120
Db :|||||
Qy 61 KGVLKAVENNEEIGDELADDDORLIDEMIKLDTANKSRLGANAILGVSMVAKA 120
Db :|||||
Qy 121 ADSAGLPLFRYIGGPNHVLVPMNIIINGGAHDSGVVDVQEFMTAPIGAETFSALRN 180
Db :|||||
Qy 121 ADSAGLPLFRYIGGPNHVLVPMNIIINGGAHDSGVVDVQEFMTAPIGAETFSALRN 180
Db :|||||
Qy 181 GAETVHALKSVIKKGLSTGLDGGFAPSVGSTREALDLIVKAIKAGTTPGKDIALAL 240
Db :|||||
Qy 181 GAETVHALKSVIKKGLSTGLDGGFAPSVGSTREALDLIVKAIKAGTTPGKDIALAL 240
Db :|||||
Qy 241 DVASSEFFKDGTYHFEGGQHSAAEMANVYAEVLDVADYPIVSIEDPQDDWEGYTNLTATI 300
Db :|||||
Qy 241 DVASSEFFKDGTYHFEGGQHSAAEMANVYAEVLDVADYPIVSIEDPQDDWEGYTNLTATI 300
Db :|||||
Qy 301 GDKVQIVGDDFFVTNPERLKEGIAKKAANSILVKVNIQIGTLTTFDAVMAHRAGYTSM 360
Db :|||||
Qy 361 SHRSGETEDTTIADLAVALNCGQIKTGAPARSDRVAKYNOLLRIEQLLDGAGVYAGRSF 420
Db :|||||
Qy 361 SHRSGETEDTTIADLAVALNCGQIKTGAPARSDRVAKYNOLLRIEQLLDGAGVYAGRSF 420
Db :|||||
Qy 421 PRFQG 425
Db :|||||
Qy 421 PRFQG 425
Db :|||||
RESULT 3
ABP97001
ID ABP97001 standard; protein; 425 AA.
XX
AC ABP97001;
XX
DT 17-JUN-2003 (first entry)
XX
XX Corynebacterium glutamicum eno protein SEQ ID NO:2.
XX
XX Fine chemical; Coryneform bacteria; Escherichia coli; microorganism;
XX genetically modified microorganism; metabolic; biosynthesis; amino acid;
XX vitamin; nucleoside; nucleotide; pigment; protein; human medicine;
XX pharmaceutical; food; animal feeding; eno.
XX
XX Corynebacterium glutamicum.

SQ Sequence 425 AA;
Query Match 99.8%; Score 2151; DB 4; Length 425;
Best Local Similarity 99.8%; Pred. No. 2.6e-169;
Matches 424; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 1 VAEIMHVFAREILDSRGNTVEAEVFLDDGSHGVAGVPSGASTGVHEAHLRDGDRYL 60
Db :|||||
Qy 1 VAEIMHVFAREILDSRGNTVEAEVFLDDGSHGVAGVPSGASTGVHEAHLRDGDRYL 60
Db :|||||
Qy 61 KGVLKAVENNEEIGDELADDDORLIDEMIKLDTANKSRLGANAILGVSMVAKA 120
Db :|||||
Qy 61 KGVLKAVENNEEIGDELADDDORLIDEMIKLDTANKSRLGANAILGVSMVAKA 120
Db :|||||
Qy 121 ADSAGLPLFRYIGGPNHVLVPMNIIINGGAHDSGVVDVQEFMTAPIGAETFSALRN 180
Db :|||||
Qy 121 ADSAGLPLFRYIGGPNHVLVPMNIIINGGAHDSGVVDVQEFMTAPIGAETFSALRN 180
Db :|||||
Qy 181 GAETVHALKSVIKKGLSTGLDGGFAPSVGSTREALDLIVKAIKAGTTPGKDIALAL 240
Db :|||||
Qy 181 GAETVHALKSVIKKGLSTGLDGGFAPSVGSTREALDLIVKAIKAGTTPGKDIALAL 240
Db :|||||
Qy 241 DVASSEFFKDGTYHFEGGQHSAAEMANVYAEVLDVADYPIVSIEDPQDDWEGYTNLTATI 300
Db :|||||
Qy 241 DVASSEFFKDGTYHFEGGQHSAAEMANVYAEVLDVADYPIVSIEDPQDDWEGYTNLTATI 300
Db :|||||
Qy 301 GDKVQIVGDDFFVTNPERLKEGIAKKAANSILVKVNIQIGTLTTFDAVMAHRAGYTSM 360
Db :|||||
Qy 301 GDKVQIVGDDFFVTNPERLKEGIAKKAANSILVKVNIQIGTLTTFDAVMAHRAGYTSM 360
Db :|||||
Qy 361 SHRSGETEDTTIADLAVALNCGQIKTGAPARSDRVAKYNOLLRIEQLLDGAGVYAGRSF 420
Db :|||||
Qy 361 SHRSGETEDTTIADLAVALNCGQIKTGAPARSDRVAKYNOLLRIEQLLDGAGVYAGRSF 420
Db :|||||
Qy 421 PRFQG 425
Db :|||||
Qy 421 PRFQG 425
Db :|||||
RESULT 2
AAG90831
ID AAG90831 standard; protein; 425 AA.
XX
AC AAG90831;
XX
DT 26-SEP-2001 (first entry)
XX
XX C glutamicum protein fragment SEQ ID NO: 4585.
XX
XX Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
XX organic acid synthesis.
XX
XX Corynebacterium glutamicum.
XX
XX EP1108790-A2.
XX
XX 20-JUN-2001.
XX
XX 18-DEC-2000; 2000EP-00127688.
XX
XX 16-DEC-1999; 99JP-00377484.
XX
XX 07-APR-2000; 2000JP-00159162.
XX
XX 03-AUG-2000; 2000TP-00280988.
XX
XX (KYOWA) KYOWA HAKKO KOGYO KK.
XX
XX Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
XX Tateishi N, Senoh A, Ikeda M, Ozaki A;
XX WPI: 2001-376931/40.
XX
XX N-PSDB; AAH66050.
XX
XX Novel polynucleotides derived from Coryneform bacteria, for identifying

•

PT carbohydrates or enzymes.
 XX Claim 20; Page 244-245; 1246pp; English.
 XX AAF71360 to AAF71750 encode the Corynebacterium glutamicum sugar
 CC metabolism and oxidative phosphorylation (SMP) proteins given in AAB79243
 CC to AAB 79633 which are involved in carbon metabolism and energy
 CC production. The C. glutamicum SMP gene can be used in vectors (II) for
 CC expression in host cells and production or modulation of production of
 CC fine chemicals, such as, an organic acid, a proteinogenic or
 CC nonproteinogenic amino acid (preferred), a purine or pyrimidine base, a
 CC nucleoside, a nucleotide, a lipid, a saturated or unsaturated fatty acid,
 CC a diol, a carbohydrate, an aromatic compound, a vitamin, a cofactor, a
 CC polyketide, or an enzyme. The presence of (I) or SMP proteins (III)
 CC encoded by them are used for diagnosing the presence or activity of
 CC Corynebacterium diphtheriae in a subject. (I), (II), (III) or host cells
 CC containing them are used to map genomes of organisms related to C.
 CC glutamicum, identify and localize C. glutamicum sequences of interest, in
 CC evolutionary studies, in determining SMP protein regions required for
 CC function, in modulating SMP protein activity, in modulating the
 CC metabolism of sugars, and in modulating high-energy molecule production
 CC in a cell (i.e. ATP, NADPH)
 XX
 SQ Sequence 425 AA;

Query Match 99.5%; Score 2145; DB 4; Length 425;
 Best Local Similarity 99.5%; Pred. No. 8.1e-169;
 Matches 423; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 VAEIMHVFAREILDSRGNTVEAEVFLDDGSHGVAGVPSGASTGVHEAHELRDGGDRYL 60
 DB 1 VAEIMHVFAREILDSRGNTVEAEVFLDDGSHGVAGVPSGASTGVHEAHELRDGGDRYL 60
 QY 61 KGVLKAVENNVNEIGDELADDDORLIDEAMIKLDGTANKSRGKAGNAILGVSMVAKA 120
 DB 61 KGVLKAVENNVNEIGDELADDDORLIDEAMIKLDGTANKSRGKAGNAILGVSMVAKA 120
 QY 121 AADSAGLPLFRYIGGPNNAHVLVPMNNIINGGAHDSGVVDVQEFMIAPIGAEFTFSALRN 180
 DB 121 AADSAGLPLFRYIGGPNNAHVLVPMNNIITGGAHDSGVVDVQEFMIAPIGAEFTFSALRN 180
 QY 181 GAEVYHALKSVIKKGLSTGLDEGGFAPSVGSTREALDLIVKAIKAGTTPGKDIALAL 240
 DB 181 GAEVYHALKSVIKKGLSTGLDEGGFAPSVGSTREALDLIVKAIKAGTTPGKDIALAL 240
 QY 241 DVASSFFKDGTHFEGGQHSAAEMANVYAEVLDAVPIVSIEDPLQDDWEGTNTLTATI 300
 DB 241 DVASSFFKDGTHFEGGQHSAAEMANVYAEVLDAVPIVSIEDPLQDDWEGTNTLTATI 300
 QY 301 GDKVQIVGDDFFVTNPERLKEGIAKKAANSILVKVNIQIGTLTETFDVDMHRAGYTSMW 360
 DB 301 GDKVQIVGDDFFVTNPERLKEGIAKKAANSILVKVNIQIGTLTETFDVDMHRAGYTSMW 360
 QY 361 SHRSGETEDTTIADLVALNCGQIKTGAPARSDRVAKYNNQILLRIEQLLDGAGYAGRSAP 420
 DB 361 SHRSGETEDTTIADLVALNCGQIKTGAPARSDRVAKYNNQILLRIEQLLDGAGYAGRSAP 420
 QY 421 PRFQG 425
 DB 421 PRFQG 425

RESULT 5
 ADD13323
 ID ADD13323 standard; protein; 425 AA.
 XX
 AC ADD13323;
 XX
 DT 01-JAN-2004 (first entry)
 XX C. glutamicum carbon metabolism associated protein RXA00235.
 DE carbon metabolism; energy-rich molecule; oxidative phosphorylation;
 KW

KW fine chemical; amino acid production; lysine production;
 KW nucleotide production; nucleoside production; lipid production;
 KW fatty acid production; diol production; carbohydrate production;
 KW aromatic compound production; vitamin production; co-factor production;
 KW enzyme production; food; animal feed; cosmetic; pharmaceutical.
 XX Corynebacterium glutamicum.
 OS
 XX Key Location/Qualifiers
 FH Misc-difference 223
 FT /note= "Optionally substituted with Lys"
 XX WO2003040291-A2.
 PN 15-MAY-2003.
 PD 31-OCT-2002; 2002WO-EP012135.
 XX 05-NOV-2001; 2001DE-01054270.
 PR (BADI) BASF AG.
 XX Zelder O, Pompejus M, Schroeder H, Kroeger B, Klopprogge C;
 PI Habershauer G;
 XX WPI; 2003-505068/47.
 DR N-PSDB; ADD13322.
 XX New nucleic acid encoding variant forms of proteins involved in carbon
 PT compound metabolism, useful for production of fine chemicals,
 PT specifically lysine, in microorganisms.
 XX Claim 1; SEQ ID NO 8; 259pp; German.
 XX This invention describes novel polynucleotides and polypeptides
 CC associated with the metabolism of carbon compounds and generation of
 CC energy-rich molecules by oxidative phosphorylation in Corynebacterium
 CC glutamicum. The polynucleotides of the invention are isolated from a
 CC nucleic acid library of C. glutamicum then mutated at the specified
 CC positions, cloned and expressed by standard methods. Cells, especially
 CC Corynebacterium glutamicum, containing vectors that express the
 CC polynucleotides are used for production of fine chemicals, preferably
 CC amino acids and specifically lysine, but more generally nucleotides,
 CC nucleosides, lipids, fatty acids, diols, carbohydrates, aromatic
 CC compounds, vitamins, co-factors and enzymes. These are useful in the
 CC food, animal feed, cosmetics and pharmaceutical industries.
 CC Polynucleotides, optionally as primers and probes, can also be used for
 CC identification and classification of C. glutamicum and related species,
 CC e.g. for diagnosis, for genomic mapping; functional or evolutionary
 CC studies gene manipulation, and modulation of metabolic activity. Cells
 CC that contain the products of the invention may produce fine chemicals in
 CC better yields, with higher productivity and/or more efficiently.

Sequence 425 AA;
 Query Match 99.5%; Score 2145; DB 7; Length 425;
 Best Local Similarity 99.5%; Pred. No. 8.1e-169;
 Matches 423; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 VAEIMHVFAREILDSRGNTVEAEVFLDDGSHGVAGVPSGASTGVHEAHELRDGGDRYL 60
 DB 1 VAEIMHVFAREILDSRGNTVEAEVFLDDGSHGVAGVPSGASTGVHEAHELRDGGDRYL 60
 QY 61 KGVLKAVENNVNEIGDELADDDORLIDEAMIKLDGTANKSRGKAGNAILGVSMVAKA 120
 DB 61 KGVLKAVENNVNEIGDELADDDORLIDEAMIKLDGTANKSRGKAGNAILGVSMVAKA 120
 QY 121 AADSAGLPLFRYIGGPNNAHVLVPMNNIINGGAHDSGVVDVQEFMIAPIGAEFTFSALRN 180
 DB 121 AADSAGLPLFRYIGGPNNAHVLVPMNNIITGGAHDSGVVDVQEFMIAPIGAEFTFSALRN 180
 QY 181 GAEVYHALKSVIKKGLSTGLDEGGFAPSVGSTREALDLIVKAIKAGTTPGKDIALAL 240
 DB 181 GAEVYHALKSVIKKGLSTGLDEGGFAPSVGSTREALDLIVKAIKAGTTPGKDIALAL 240

Db 181 GAETHALKSVIKKGLSTGLGDEGGFAPSVGSTREALDLIVEAEKAGFTPGKDIALAL 240
 QY 241 DVASSEFFKDGTYHFEQGHSAEMANVYAEVLVDAYPIVSIEDPLQEDDWEQYTNLTATI 300
 Db 241 DVASSEFFKDGTYHFEQGHSAEMANVYAEVLVDAYPIVSIEDPLQEDDWEQYTNLTATI 300
 QY 301 GSKVQIVGDDFFVTNPERLKEGIAKKAANSILVKVNIQIGTLTETFDVDMHRAGYTSM 360
 Db 301 GSKVQIVGDDFFVTNPERLKEGIAKKAANSILVKVNIQIGTLTETFDVDMHRAGYTSM 360
 QY 361 SHRSGETEDTTIADLAVALNCGQIKTGAPARSDRVAKYNOLLRIEQLLGDAGVYAGRSAP 420
 Db 361 SHRSGETEDTTIADLAVALNCGQIKTGAPARSDRVAKYNOLLRIEQLLGDAGVYAGRSAP 420
 QY 421 PRFOG 425
 Db 421 PRFOG 425

RESULT 6
 ABP97002
 ID ABP97002 standard; protein; 425 AA.
 XX
 AC ABP97002;
 DT 17-JUN-2003 (first entry)
 DE Modified eno S330E protein SEQ ID NO:6.
 XX

KW Fine chemical; Coryneform bacteria; Escherichia coli; microorganism;
 KW genetically modified microorganism; metabolite; biosynthesis; amino acid;
 KW vitamin; nucleoside; nucleotide; pigment; protein; human medicine;
 KW pharmaceutical; food; animal feeding; eno.

XX Corynebacterium glutamicum.
 OS Synthetic.
 OS

XX WO2003023016-A2.
 XX 20-MAR-2003.
 XX 11-SEP-2002; 2002WO-EP010174.
 XX 13-SEP-2001; 2001DE-01045043.
 XX (DEGS) DEGUSSA AG.
 FA Farwick M, Hermann T;
 FI WPI; 2003-354534/33.
 DR N-PSDB; ACC45314.

XX Microorganism useful for producing e.g. fine chemicals, has permanently
 PT altered phosphorylatability protein, such that biosynthesis of fine
 PT chemical synthesized by microorganism is increased compared to wild-type.
 XX
 PS Example 3; Page 85-86; 120pp; English.

CC The present invention describes a microorganism (I), in which the
 CC phosphorylatability of at least one protein has been permanently altered
 CC such that the biosynthesis of at least one fine chemical synthesised by
 CC the microorganism is increased compared to the wild type. Also described:
 CC (1) use of a DNA (II) sequence coding for a protein which contains a
 CC phosphorylation site, where the sequence contains such a mutation that
 CC the protein is changed in its phosphorylatability for the production of
 CC (I), or for the production of fine chemicals; and (2) a method for
 CC producing fine chemicals or metabolites comprising using (I). (I) is
 CC useful for producing fine chemicals or metabolites, such as amino acids,
 CC vitamins, nucleosides, nucleotides, pigments or proteins. The amino acids
 CC and vitamins produced using (I) can be used in human medicine, in the
 CC pharmaceutical industry, food industry and in animal feeding. (I)
 CC produces larger amount of desired fine chemical or a metabolite than the
 CC wild type. The present sequence represents a modified eno S330E from

CC Corynebacterium glutamicum, which is used in an example from the present
 CC invention
 XX
 SQ Sequence 425 AA;
 Query Match 99.5%; Score 2144; DB 6; Length 425;
 Best Local Similarity 99.3%; Pred. No. 9.8e-169;
 Matches 422; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 1 VAETMHVPAREILDSRGNPTVEAEVFLDDGSHGVAGVPSGASTGVHEAHELRDGGDRYL 60
 Db 1 MAETMHVPAREILDSRGNPTVEAEVFLDDGSHGVAGVPSGASTGVHEAHELRDGGDRYL 60
 QY 61 KGVLKAVENNVNEEIGDELGLAGLEADDQRLIDEAMTKLDGTANKSRLGANAILGVSNVAKA 120
 Db 61 KGVLKAVENNVNEEIGDELGLAGLEADDQRLIDEAMTKLDGTANKSRLGANAILGVSNVAKA 120
 QY 121 AADSAGLPLFRYIGGPNHVLVPEPMNIINGGAHADSGVDVQEFMIPIGAETFEALRN 180
 Db 121 AADSAGLPLFRYIGGPNHVLVPEPMNIINGGAHADSGVDVQEFMIPIGAETFEALRN 180
 QY 181 GAETHALKSVIKKGLSTGLGDEGGFAPSVGSTREALDLIVKAEKAGFTPGKDIALAL 240
 Db 181 GAETHALKSVIKKGLSTGLGDEGGFAPSVGSTREALDLIVKAEKAGFTPGKDIALAL 240
 QY 241 DVASSEFFKDGTYHFEQGHSAEMANVYAEVLVDAYPIVSIEDPLQEDDWEQYTNLTATI 300
 Db 241 DVASSEFFKDGTYHFEQGHSAEMANVYAEVLVDAYPIVSIEDPLQEDDWEQYTNLTATI 300
 QY 301 GSKVQIVGDDFFVTNPERLKEGIAKKAANSILVKVNIQIGTLTETFDVDMHRAGYTSM 360
 Db 301 GSKVQIVGDDFFVTNPERLKEGIAKKAANSILVKVNIQIGTLTETFDVDMHRAGYTSM 360
 QY 361 SHRSGETEDTTIADLAVALNCGQIKTGAPARSDRVAKYNOLLRIEQLLGDAGVYAGRSAP 420
 Db 361 SHRSGETEDTTIADLAVALNCGQIKTGAPARSDRVAKYNOLLRIEQLLGDAGVYAGRSAP 420
 QY 421 PRFOG 425
 Db 421 PRFOG 425

RESULT 7
 ABU25984
 ID ABU25984 standard; protein; 425 AA.
 XX
 AC ABU25984;
 DT 19-JUN-2003 (first entry)
 XX
 DE Protein encoded by Prokaryotic essential gene #11511.
 XX Antisense; prokaryotic essential gene; cell proliferation; drug design.
 XX Corynebacterium diphtheriae.
 XX WO200277183-A2.
 XX 03-OCT-2002.
 XX 21-MAR-2002; 2002WO-US009107.
 XX 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 XX (ELIT-) ELITRA PHARM INC.
 XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

Db 364 RSGETEDTTIADLAVAGSGQIKTGAPARSAKYNQLLRIBEALGDAARYAGDLAFPR 423
 Qy 423 F 423
 Db 424 F 424

RESULT 10
 ABU36566
 ID ABU36566 standard; protein; 429 AA.
 XX AC ABU36566;
 XX 19-JUN-2003 (first entry)
 XX DE Protein encoded by Prokaryotic essential gene #22093.
 XX KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
 XX OS Mycobacterium tuberculosis.
 XX PN WO200277183-A2.
 XX PD 03-OCT-2002.
 XX PF 21-MAR-2002; 2002WO-US009107.
 XX PR 21-MAR-2001; 2001US-00815242.
 XX PR 06-SEP-2001; 2001US-00948993.
 XX PR 25-OCT-2001; 2001US-0342923P.
 XX PR 08-FEB-2002; 2002US-00072851.
 XX PR 06-MAR-2002; 2002US-0362699P.
 XX PA (ELIT-) ELITRA PHARM INC.
 XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 XX PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX WPI; 2003-029926/02.
 XX DR N-PSDB; ACA40436.
 XX PT New antisense nucleic acids, useful for identifying proteins or screening
 XX PT for homologous nucleic acids required for cellular proliferation to
 XX PT isolate candidate molecules for rational drug discovery programs.
 XX PS Claim 25; SEQ ID NO 64490; 1766pp; English.
 XX CC The invention relates to an isolated nucleic acid comprising any one of
 XX CC the 6213 antisense sequences given in the specification where expression
 XX CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 XX CC (1) a vector comprising a promoter operably linked to the nucleic acid
 XX CC encoding a polypeptide whose expression is inhibited by the antisense
 XX CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 XX CC polypeptide or its fragment whose expression is inhibited by the
 XX CC antisense nucleic acid; (4) an antibody capable of specifically binding
 XX CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 XX CC proliferation or the activity of a gene in an operon required for
 XX CC proliferation; (7) identifying a compound that influences the activity of
 XX CC the gene product or that has an activity against a biological pathway
 XX CC required for proliferation, or that inhibits cellular proliferation; (8)
 XX CC identifying a gene required for cellular proliferation or the biological
 XX CC pathway in which a proliferation-required gene or its gene product lies
 XX CC or a gene on which the test compound that inhibits proliferation of an
 XX CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 XX CC compound's activity; (11) a culture comprising strains in which the gene
 XX CC product is overexpressed or underexpressed; (12) determining the extent
 XX CC to which each of the strains is present in a culture or collection of
 XX CC strains; or (13) identifying the target of a compound that inhibits the
 XX CC proliferation of an organism. The antisense nucleic acids are useful for
 XX CC identifying proteins or screening for homologous nucleic acids required
 XX CC for cellular proliferation to isolate candidate molecules for rational
 XX CC drug discovery programs, or for screening homologous nucleic acids
 XX CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC K. pneumoniae or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

Query Match 72.1%; Score 1553.5; DB 6; Length 429;
 Best Local Similarity 73.2%; Pred. No. 7.6e-120;
 Matches 308; Conservative 42; Mismatches 70; Indels 1; Gaps 1;
 SQ Sequence 429 AA;

Qy 4 IMHVFAREILDSRGNTVEAEVFLDDGSHGVACVPSGASTGVHAEHLRDPGRYLGKV 63
 Db 4 IEQVRAREILDSRGNTVEAEVFLDDGSHGVACVPSGASTGVHAEHLRDPGRYLGKV 63
 Qy 64 LKAVENVNEIGDELACLEADDDORLIDEAMIKLDGTANKSRIGCANAILGVSMVAKAAD 123
 Db 64 QKAVQAVLDDEIGPAVIGLNAADDORLVDQALVDLDGTPDKSRIGCANAILGVSLAVAKAAD 123
 Qy 124 SAGLPLFRYIGGPNNAHVLVPMNNIINGGAHDSGVDOEFMIAPIGAETFSBALRNGAE 183
 Db 124 SAGLPLFRYIGGPNNAHVLVPMNNIINGGAHDSGVDOEFMIAPIGAETFSBALRNGAE 183
 Qy 184 VHALKSVLKEGLSTGLDGGFAPSVGSTREALDLIVKAEKAGFTPGKDIALADVA 243
 Db 184 VHALKSVLKEGLSTGLDGGFAPSVGSTREALDLIVKAEKAGFTPGKDIALADVA 243
 Qy 244 SSEFFKDGDT-YHPEGGQHSNAEMANYAEIVDAVPIVSTEDPQEDDWEQYTNLTATIGD 302
 Db 244 ATEFFTDGTYGVEPTGTRTADQMTFYGALLGAYPLVSIEDPLSEDDWDGWAALTASIGD 303
 Qy 303 KVOIVGDDPFVTNPERLKEGIAKKAANSIIKVNQIGTLTETDPAVDMHRAGYTSMMSH 362
 Db 304 RVQIVGDDPFVTNPERLESGIERGVANALLKVNQIGTLTETDPAVDMHRAGYTSMMSH 363
 Qy 363 RSGETEDTTIADLAVAGSGQIKTGAPARSAKYNQLLRIBEALGDAARYAGDLAFPR 422
 Db 364 RSGETEDTTIADLAVAGSGQIKTGAPARSAKYNQLLRIBEALGDAARYAGDLAFPR 423
 Qy 423 F 423
 Db 424 F 424

RESULT 11
 ABU34772
 ID ABU34772 standard; protein; 439 AA.
 XX AC ABU34772;
 XX 19-JUN-2003 (first entry)
 XX DE Protein encoded by Prokaryotic essential gene #20299.
 XX KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
 XX OS Mycobacterium bovis.
 XX PN WO200277183-A2.
 XX PD 03-OCT-2002.
 XX PF 21-MAR-2002; 2002WO-US009107.
 XX PR 21-MAR-2001; 2001US-00815242.
 XX PR 06-SEP-2001; 2001US-00948993.
 XX PR 25-OCT-2001; 2001US-0342923P.
 XX PR 08-FEB-2002; 2002US-00072851.
 XX PR 06-MAR-2002; 2002US-0362699P.
 XX PA (ELIT-) ELITRA PHARM INC.

PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX WPI: 2003-029926/02.
 DR N-PSDB; AAC38642.
 XX
 PT New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX
 PS Claim 25; SEQ ID NO 62696; 1766pp; English.
 XX
 CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 439 AA;

Query Match 72.1%; Score 1553.5; DB 6; Length 439;
 Best Local Similarity 73.2%; Pred. No. 7,9e-120;
 Matches 308; Conservative 42; Mismatches 70; Indels 1; Gaps 1;

QY 4 IMHVPAREILDSRGNTVEAEVFLDDGSHGVAGVPSGASTGVHNEHURDGGDRLGKGV 63
 DB 14 IEQVGAREILDSRGNTVEAEVFLDDGSHGVAGVPSGASTGVHNEHURDGGDRLGKGV 73
 QY 64 LKAVENNEEIGDELAGEADDORLIDRAMIKLDTANKSRIGANAIIGVSMVAKAARD 123
 DB 74 QKAVQVLEIDEGPAGVIGLNADDORLIDRAMIKLDTANKSRIGANAIIGVSMVAKAARD 133
 QY 124 SAGLPLFRYIGPNNAHVLFPVPMNIIINGGAHDSGVYQVEFMFIAPIGATFSEALRNGAE 183
 DB 134 SAEPLFLFRYIGPNNAHVLFPVPMNIIINGGAHDSGVYQVEFMFIAPIGATFSEALRNGAE 193
 QY 184 VTHALKSVIKKGLSTGLDEGGFAPSVGSTREALDLIVKAEKAGFTPGKDIALDLVA 243
 DB 194 VTHALKSVIKKGLSTGLDEGGFAPSVGSTREALDLIVKAEKAGFTPGKDIALDLVA 253
 QY 244 SSEFFKCT-YHFEQGHQSAEMANVYAEVDAYPIVSIEDPLOBDDWEGTNTATIGD 302
 DB 254 ATEFTDGTGVYFEGTETADQMTBFYAGLGGAYPLVSIEDPLOBDDWEGTNTATIGD 313
 QY 303 KVQIVGDDFFVTNPERLKEGIAKKAANSILVKVNOIGTLTETFDVADMAHRAGYTSMHSH 362
 DB 314 RVQIVGDDIFVTNPERLKEGIERGVANALLVKVNOIGTLTETLDVTLAHGGYRTMISH 373

QY 363 RSGETEDTTIADLAVLNCGQIKTGAPARSDRVAKYNOLLRIEQILGDAGVYAGSAPPR 422
 DB 374 RSGETEDTTIADLAVLNCGQIKTGAPARSDRVAKYNOLLRIEQILGDAGVYAGSAPPR 433
 QY 423 F 423
 DB 434 F 434
 RESULT 12
 AAU54498
 ID AAU54498 standard; protein; 455 AA.
 AC AAU54498;
 XX
 DT 27-FEB-2002 (first entry)
 DE Propionibacterium acnes immunogenic protein #15394.
 KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
 KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 KW dermatological; osteopathic; neuroprotectant.
 OS Propionibacterium acnes.
 XX WO200181581-A2.
 XX
 PD 01-NOV-2001.
 XX
 PF 20-APR-2001; 2001WO-US012865.
 XX
 PR 21-APR-2000; 2000US-0199047P.
 PR 02-JUN-2000; 2000US-0208841P.
 PR 07-JUL-2000; 2000US-0216747P.
 XX
 PA (CORI-) CORIXA CORP.
 PI Skeiky YAM, Persing DH, Mitcham JL, Wang SS, Bhatia A;
 PI L'maisonmeuve J, Zhang Y, Jen S, Carter D;
 XX
 DR WPI: 2001-616774/71.
 DR N-PSDB; AAS59564.
 XX
 PT Propionibacterium acnes polypeptides and nucleic acids useful for
 PT vaccinating against and diagnosing infections, especially useful for
 PT treating acne vulgaris.
 XX
 PS Claim 6; SEQ ID NO 15693; 1069pp; English.
 XX
 CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
 CC polypeptides. The proteins and their associated DNA sequences are used in
 CC the treatment, prevention and diagnosis of medical conditions caused by
 CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
 CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
 CC P. acnes is also involved in infections of bone, joints and the central
 CC nervous system, however it is particularly involved in the inflammatory
 CC lesions associated with acne vulgaris. A method for detecting the
 CC presence or absence of P. acnes in a patient comprises contacting a
 CC sample with a binding agent that binds to the proteins of the invention
 CC and determining the amount of bound protein in the sample. The
 CC polypeptides may be used as antigens in the production of antibodies
 CC specific for P. acnes proteins. These antibodies can be used to
 CC downregulate expression and activity of P. acnes polypeptides and
 CC therefore treat P. acnes infections. The antibodies may also be used as
 CC diagnostic agents for determining P. acnes presence, for example, by
 CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
 CC this patent did not form part of the printed specification, but was
 CC obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 455 AA;

ABU35758;
19-JUN-2003 (first entry)
Protein encoded by Prokaryotic essential gene #2185.
Antisense; prokaryotic essential gene; cell proliferation; drug design.
Mycobacterium leprae.
W0200277183-A2.
03-OCT-2002.
21-MAR-2002; 2002WO-US009107.
21-MAR-2001; 2001US-00815242.
06-SEP-2001; 2001US-00948993.
25-OCT-2001; 2001US-0342923P.
08-FEB-2002; 2002US-00072851.
06-MAR-2002; 2002US-0362699P.
(ELIT-) ELITRA PHARM INC.
Wang L, Zamudio C, Malone C, Haseibeck R, Ohlsen KL, Zyskind JW;
Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
WPI; 2003-0299926/02.
N-PSDB; ACA39628.
New antisense nucleic acids, useful for identifying proteins or screening
for homologous nucleic acids required for cellular proliferation to
isolate candidate molecules for rational drug discovery programs.
Claim 25; SEQ ID NO 63682; 1765pp; English.
The invention relates to an isolated nucleic acid comprising any one of
the 6213 antisense sequences given in the specification where expression
of the nucleic acid inhibits proliferation of a cell. Also included are:
(1) a vector comprising a promoter operably linked to the nucleic acid
encoding a polypeptide whose expression is inhibited by the antisense
nucleic acid; (2) a host cell containing the vector; (3) an isolated
polypeptide or its fragment whose expression is inhibited by the
antisense nucleic acid; (4) an antibody capable of specifically binding
the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
proliferation or the activity of a gene in an operon required for
proliferation; (7) identifying a compound that influences the activity of
the gene product or that has an activity against a biological pathway
required for proliferation, or that inhibits cellular proliferation; (8)
identifying a gene required for cellular proliferation or the biological
pathway in which a proliferation-required gene or its gene product lies
or a gene on which the test compound that inhibits proliferation of an
organism acts; (9) manufacturing an antibiotic; (10) profiling a
compound's activity; (11) a culture comprising strains in which the gene
product is overexpressed or underexpressed; (12) determining the extent
to which each of the strains is present in a culture or collection of
strains; or (13) identifying the target of a compound that inhibits the
proliferation of an organism. The antisense nucleic acids are useful for
identifying proteins or screening for homologous nucleic acids required
for cellular proliferation to isolate candidate molecules for rational
drug discovery programs, or for screening homologous nucleic acids
required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
K. pneumoniae or *P. aeruginosa*. The present sequence is encoded by one of
the target prokaryotic essential genes. Note: The sequence is encoded by one of
patent did not form part of the printed specification, but was obtained
in electronic format directly from WIPO at
ftp.wipo.int/pub/published_pct_sequences
Sequence 447 AA;
Query Match 69.7%; Score 1502.5; DB 6; Length 447;
Best Local Similarity 70.8%; Pred NO. 1.3e-115;

PT New recombinant DNA construct comprising a promoter positioned to provide
 PT for expression of a polynucleotide encoding a polypeptide from a
 XX microbial source, useful for producing plants with improved properties.
 XX
 PS Claim 1; SEQ ID NO 11696; 122pp; English.
 XX
 CC The invention relates to a recombinant DNA construct comprising a
 CC promoter functional in a plant cell, where the promoter is positioned to
 CC provide for expression of a polynucleotide encoding a polypeptide from a
 CC microbial source. The invention also relates to a transformed plant
 CC comprising the recombinant DNA construct and a method of producing a
 CC transformed plant having an improved property. The plant is a crop plant
 CC such as maize or soybean. The method of producing a transformed plant
 CC having an improved property comprises transforming a plant with the
 CC recombinant DNA construct and growing the transformed plant, where the
 CC polynucleotide or polypeptide is useful for improving plant properties.
 CC The recombinant DNA construct is useful for producing plants with
 CC improved plant properties, e.g. improved cold, heat or drought tolerance,
 CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
 CC increased resistance to plant disease, better growth rate by modification
 CC of the cell cycle pathway with plant growth regulators, increased rate of
 CC homologous recombination, modified seed oil or protein yield and/or
 CC content, improved yield by modification of carbohydrate, nitrogen or
 CC phosphorus use and/or uptake, by modification of photosynthesis or by
 CC providing improved plant growth and development under at least one stress
 CC condition, improved lignin production or improved galactomannan
 CC production. This sequence represents a bacterial polypeptide used in the
 CC scope of the invention. Note: The sequence data for this patent did not
 CC form part of the printed specification but was obtained in electronic
 CC format from USPTO at seqdata.uspto.gov/sequence.html.
 XX
 XX Sequence 423 AA;

Query Match	66.0%	Score 1422;	DB 8;	Length 423;
Best Local Similarity	66.2%	Pred. No. 5.8e-109;		
Matches	276;	Conservative 58;	Mismatches 81;	Indels 2; Gaps 1;

Qy	4	IMHVFAEILDSRGNTVEAEVFLDDGSHGVAGVPSGASTGVHAEHLRDGGDRYLKGV	63
Db	4	ITDIAREILDSRGNTVEVDVYLEDGSMGRAAVPSGASTGAHEAVELRDGGKRYLKG	63
Qy	64	LKAVENVEETIGDELADDOBLIDEAMTKLDGTANKSRLGANAILGVSMVAKAAAD	123
Db	64	EKAVEAVNTEIFDAIGDFDAENQIQIDQMIALDGTENKSLGANAILGVSLAKAAE	123
Qy	124	SAGLPFRYIGGPNNAHVLVPMNIIINGGAHDSVDVQEFMIAPIGATFSEALRNGAE	183
Db	124	ASGLPLYRYVGGPNNAHLLVPMNIIINGGAHADNPIDQEFMILPVGAENIRRAVRMGSE	183
Qy	184	VYHALKSVIKEGLSTGLDEGGFAPSVGSTREALDLIVKAEKAGFTPGKDIALDLVA	243
Db	184	VFHTLKKELSAQGHNTVNGDEGGFAPGLSSAPALDFIMKSIKAGYRPGEDMYVGLDCA	243
Qy	244	SSSEFFKDGTYHFEGGQHS--AAEMANVYAEIVDAPIVTSIEDPLQEDDEGYTNLTATIG	301
Db	244	STFEFKDGKYLEGEGHTLEPGAWAYLAELVNKYPISVEDGHAEDDEGWTLTDLVG	303
Qy	302	DKQIVGDDFFVTNPERLKEGIAKKAANSILVKVNQIGTTLTETFDVDMAHRAGYTSMMS	361
Db	304	NKCQLVGDDLVFTNSARLRDGIKMGVANSILVKVNQIGSLSETLDAVETAKAGYTAAMS	363
Qy	362	HRSGETEDTTIADLAVNLNGQIKTGAPASDRVAKNQLRIEQLLDGAGVYAGRS	418
Db	364	HRSGETEDTTIADLAVNCGQINTGSLARSRLAKYNQLRIEEMILGPOAYAGRS	420

Search completed: September 29, 2005, 12:10:39
 Job time : 94.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 28, 2005, 16:08:34 ; Search time 275.5 Seconds
(without alignments)
9372.213 Million cell updates/sec

Title: US-10-728-947-1

Perfect score: 1578

Sequence: 1 ggctgggagatgggtagtt.....ctcaagcagggaacgtgctt 1578

Scoring table:

IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA.*
1: /cgn2_6/ptodata/1/ina/5A COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PTUS COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1578	100.0	1578	4	US-09-860-768-1
2	1576.4	99.9	1578	4	US-09-860-768-3
3	632.6	40.1	4403765	3	US-09-103-840A-2
4	632.6	40.1	4411529	3	US-09-103-840A-1
5	552	35.0	1299	4	US-09-902-540-4780
6	552	35.0	27903	4	US-09-902-540-1235
7	533.6	33.8	1335	4	US-09-252-991A-4449
8	533	33.8	1347	4	US-09-252-991A-4158
9	481.2	30.5	1380	4	US-09-489-039A-2890
10	449.2	28.5	2363	4	US-09-818-780-22
11	442.4	28.0	3545	2	US-08-781-802-9
12	433.2	27.5	1305	4	US-09-583-110-1156
13	432	27.4	11384	3	US-08-961-527-45
14	429.6	27.2	1196	4	US-09-107-532A-156
15	420.2	26.6	1233	4	US-09-489-039A-2858
16	415.6	26.3	1511	3	US-08-847-065-18
17	415.6	26.3	1511	4	US-09-829-382-18
18	391.8	24.8	3513	3	US-09-058-260-1
19	391.8	24.8	4315	2	US-08-781-802-1
20	391.8	24.8	4315	3	US-08-694-078-1
21	391.2	24.8	1830121	4	US-09-557-884-1
22	391.2	24.8	1830121	4	US-09-543-990A-1
23	385.4	24.4	1371	4	US-09-543-681A-1918
24	383.2	24.3	1224	4	US-09-107-433-2383
25	374.8	23.8	1335	4	US-09-328-352-3143
26	367.2	23.3	1338	4	US-09-540-236-434
27	367.2	23.3	92407	4	US-09-596-002-36

28	359.6	22.8	1460	4	US-09-830-217-21	Sequence 21, Appl
29	359.6	22.8	1460	4	US-10-278-946-21	Sequence 21, Appl
30	349.6	22.2	1305	4	US-09-710-279-2321	Sequence 221, Ap
31	349.6	22.2	1323	3	US-09-134-001C-2804	Sequence 2804, Ap
32	349.6	22.2	3830	4	US-09-710-279-3773	Sequence 3773, Ap
33	344.6	21.8	1539	4	US-08-956-171E-207	Sequence 207, App
34	344.6	21.8	1539	4	US-08-781-986A-207	Sequence 207, App
35	340.6	21.6	1230230	4	US-09-438-185A-1	Sequence 1, Appli
36	330.6	21.0	1664976	4	US-08-916-421B-1	Sequence 1, Appli
37	330.6	21.0	1664976	4	US-09-692-570-1	Sequence 1, Appli
38	330.2	20.9	1230025	4	US-09-198-452A-1	Sequence 1, Appli
39	309.6	19.6	1314	4	US-09-614-221A-509	Sequence 509, App
40	306.4	19.4	1314	4	US-09-614-221A-466	Sequence 466, App
41	303.2	19.2	1884	4	US-09-566-921-38	Sequence 38, Appl
42	302.6	19.2	934	3	US-08-961-083-183	Sequence 183, App
43	302.6	19.2	934	4	US-09-536-784-183	Sequence 183, App
44	301.2	19.1	2422	4	US-09-949-016-282	Sequence 282, App
45	301.2	19.1	2422	4	US-09-949-016-2488	Sequence 2488, Ap

ALIGNMENTS

RESULT 1
US-09-860-768-1
; Sequence 1, Application US/09860768
; Patent No. 6713289
; GENERAL INFORMATION:
; APPLICANT: Mockel, Bettina
; APPLICANT: Pfeifferle, Walter
; APPLICANT: Hermann, Thomas
; APPLICANT: Pohler, Alfred
; APPLICANT: Kalinowski, Jörn
; APPLICANT: Bathe, Brigitte
; TITLE OF INVENTION: New Nucleotide Sequences that Code for the Eno Gene
; FILE REFERENCE: 21123/278404
; CURRENT APPLICATION NUMBER: US/09/860.768
; CURRENT FILING DATE: 2001-05-21
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 1578
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
US-09-860-768-1

Query Match	100.0%;	Score 1578;	DB 4;	Length 1578;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1578;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	GGCTGGGGATATGGGTAGTTTTCGCCCACTAATTTCAACTGATTCGCTCATCGAAACAGA	60	
Db	1	GGCTGGGGATATGGGTAGTTTTCGCCCACTAATTTCAACTGATTCGCTCATCGAAACAGA	60	
Qy	61	TTGCTGCAACAATTTGGGTAGACGTGATTGAAGACATTTGATCAGTGAATAATTCTAG	120	
Db	61	TTGCTGCAACAATTTGGGTAGACGTGATTGAAGACATTTGATCAGTGAATAATTCTAG	120	
Qy	121	TTAGTCTCCCAAGTTGGATAGGAGGCCACAGTGGTGAATATCATGACAGTATTCGCTGC	180	
Db	121	TTAGTCTCCCAAGTTGGATAGGAGGCCACAGTGGTGAATATCATGACAGTATTCGCTGC	180	
Qy	181	GAATTTCTCCACCTCCCGCGGTAAACCAACCGTCGAGCAGAGGTTTCTCGATGACGGT	240	
Db	181	GAATTTCTCCACCTCCCGCGGTAAACCAACCGTCGAGCAGAGGTTTCTCGATGACGGT	240	
Qy	241	TCCACCGGTTCGCGAGGTTCATCCGCGGCATCCACCGCGCTCCACGAGGCTCATGAG	300	
Db	241	TCCACCGGTTCGCGAGGTTCATCCGCGGCATCCACCGCGCTCCACGAGGCTCATGAG	300	
Qy	301	CTGCGTGACGGTGGCGATTCGCTACCTGGGCAAGGGCGTTTGAAGGAGGAGTTGAAAACGTC	360	
Db	301	CTGCGTGACGGTGGCGATTCGCTACCTGGGCAAGGGCGTTTGAAGGAGGAGTTGAAAACGTC	360	

```
QY 361 AACGAGAAATCGCGAGAGCTCGCTGGCTAGAGCTGACGATCAGCGCCTCATCGAC 420
Db 361 AACGAGAAATCGCGAGAGCTCGCTGGCTAGAGCTGACGATCAGCGCCTCATCGAC 420
QY 421 GAAGCAATGATCAAGCTTGTATGGCAGCGCAACAAAGTCCCGCTCGGTGCAAAACGCAATC 480
Db 421 GAAGCAATGATCAAGCTTGTATGGCAGCGCAACAAAGTCCCGCTCGGTGCAAAACGCAATC 480
QY 481 CTTGGTGTTCCTATGGCTGTGTCGAAAGCTGCTGCTGCTGATTCGCGAGGCTCCCACTGTTT 540
Db 481 CTTGGTGTTCCTATGGCTGTGTCGAAAGCTGCTGCTGCTGATTCGCGAGGCTCCCACTGTTT 540
QY 541 CGCTACATCGGTGAGCAACAAAGCAACAGCTTCTCCAGTTCCATGATGAACATCATCAAC 600
Db 541 CGCTACATCGGTGAGCAACAAAGCAACAGCTTCTCCAGTTCCATGATGAACATCATCAAC 600
QY 601 GGTGGCGCTCAGCTGACTCGGTGTTGAGCTTCAGGAAATTCATGATCGCTCCAATCGGT 660
Db 601 GGTGGCGCTCAGCTGACTCGGTGTTGAGCTTCAGGAAATTCATGATCGCTCCAATCGGT 660
QY 661 GCAGAGACCTTCTCTGAGGCTCTCCGCAACGGCGGAGGTCTACCAACGCACTGAAGTCC 720
Db 661 GCAGAGACCTTCTCTGAGGCTCTCCGCAACGGCGGAGGTCTACCAACGCACTGAAGTCC 720
QY 721 GTCATCAAGNAAGGCGCTGTCCACCGGACTTGGCGATGAGGGCGCTTCGCTCCTTCC 780
Db 721 GTCATCAAGNAAGGCGCTGTCCACCGGACTTGGCGATGAGGGCGCTTCGCTCCTTCC 780
QY 781 GTGGGCTCCACCGCTGAGGCTCTTGAACCTTATCGTTCAGGCAATCGAAGAGGCTGGCTTC 840
Db 781 GTGGGCTCCACCGCTGAGGCTCTTGAACCTTATCGTTCAGGCAATCGAAGAGGCTGGCTTC 840
QY 841 ACCCAGGCAAGGACATCGCTCTTGTCTGAGAGCTTCTCTCTGAGTTCTTCAAGGAC 900
Db 841 ACCCAGGCAAGGACATCGCTCTTGTCTGAGAGCTTCTCTCTGAGTTCTTCAAGGAC 900
QY 901 GGACCTTACCACTTCGAAGTGGCGAGCACTCCCGAGCTGAGATGGCAAGTTAGCT 960
Db 901 GGACCTTACCACTTCGAAGTGGCGAGCACTCCCGAGCTGAGATGGCAAGTTAGCT 960
QY 961 GAGCTCGTTGACGGTACCCCAATCGTCTCCATCGAGGACCCACCTGCAGGAAGATGACTGG 1020
Db 961 GAGCTCGTTGACGGTACCCCAATCGTCTCCATCGAGGACCCACCTGCAGGAAGATGACTGG 1020
QY 1021 GAGGTTTACCAACCTCACCGCAACCAATCGGCGACAAGGTTCAAGATCGTTGGCGAGC 1080
Db 1021 GAGGTTTACCAACCTCACCGCAACCAATCGGCGACAAGGTTCAAGATCGTTGGCGAGC 1080
QY 1081 TTCTTCGTCAACCAACCTCAGGCTTGAAGAGGGCATCGCTAAGAGGCTGCGCAATCC 1140
Db 1081 TTCTTCGTCAACCAACCTCAGGCTTGAAGAGGGCATCGCTAAGAGGCTGCGCAATCC 1140
QY 1141 ATCTCGTTAAGGTGAACAGATCGGTACCTCAGGAGCTTCGAGCGCTGTCGACATG 1200
Db 1141 ATCTCGTTAAGGTGAACAGATCGGTACCTCAGGAGCTTCGAGCGCTGTCGACATG 1200
QY 1201 GCTCACCGCGAGGCTACACCTTCCATGATGTCACACCGTTCGCGTGAGACCGAGGACAC 1260
Db 1201 GCTCACCGCGAGGCTACACCTTCCATGATGTCACACCGTTCGCGTGAGACCGAGGACAC 1260
QY 1261 ACCATTGCTGACCTCGAGTTGCACTCAACTGTGGCCAGATCAAGCTGGTCTCCAGCA 1320
Db 1261 ACCATTGCTGACCTCGAGTTGCACTCAACTGTGGCCAGATCAAGCTGGTCTCCAGCA 1320
QY 1321 GGTTCGACCGGTGTCGCAAGTACAAACAGCTTCTCCGCACTCGAGCAGCTGCTTGGCGAC 1380
Db 1321 GGTTCGACCGGTGTCGCAAGTACAAACAGCTTCTCCGCACTCGAGCAGCTGCTTGGCGAC 1380
QY 1381 GCCGCGCTTACGAGGTGCGAGCGCATTCGCCACGCTTTTCAGGGCTAAATAAAGCGCTT 1440
Db 1381 GCCGCGCTTACGAGGTGCGAGCGCATTCGCCACGCTTTTCAGGGCTAAATAAAGCGCTT 1440
```

```
QY 1441 TTCGAGCCCGGTAACTCAAGTTGCCGGGCGTCTGTTGCTTACTACTGTTACTGTTGT 1500
Db 1441 TTCGAGCCCGGTAACTCAAGTTGCCGGGCGTCTGTTGCTTACTACTGTTACTGTTGT 1500
QY 1501 GACTATGATCGAGGATTATGGCAAGCAAGAGAAAACTCATAAAGCCCTGTTCTGTTCT 1560
Db 1501 GACTATGATCGAGGATTATGGCAAGCAAGAGAAAACTCATAAAGCCCTGTTCTGTTCT 1560
QY 1561 CAAGCAGGAAACGTGCTT 1578
Db 1561 CAAGCAGGAAACGTGCTT 1578

RESULT 2
US-09-860-768-3
; Sequence 3, Application US/09860768
; Patent No. 6713289
; GENERAL INFORMATION:
; APPLICANT: Mockel, Bettina
; APPLICANT: Pfeifferle, Walter
; APPLICANT: Hermann, Thomas
; APPLICANT: Pohler, Alfred
; APPLICANT: Kalinowski, Jorn
; APPLICANT: Bathe, Brigitte
; TITLE OF INVENTION: New Nucleotide Sequences that Code for the Eno Gene
; FILE REFERENCE: 21123/278404
; CURRENT APPLICATION NUMBER: US/09/860,768
; CURRENT FILING DATE: 2001-05-21
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 1578
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
US-09-860-768-3

Query Match 99.9%; Score 1576.4; DB 4; Length 1578;
Best Local Similarity 99.9%; Pred.No. 0;
Matches 1577; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGCTGGGGATATCGGTAGTTTTCGCCACTAATTTCAACTGATTCCTCATCGAAACAAGA 60
Db 1 GGCTGGGGATATCGGTAGTTTTCGCCACTAATTTCAACTGATTCCTCATCGAAACAAGA 60
QY 61 TTCTGTCACAAATTTGGGTAGACGTGATTGAAGACATTTTGATCAGTGAATATTTCTAG 120
Db 61 TTCTGTCACAAATTTGGGTAGACGTGATTGAAGACATTTTGATCAGTGAATATTTCTAG 120
QY 121 TTAGCTCCCAAGTTGGCATAGGAGGCCACAGTGGCTGAAATCATGCACTATTCGCTCGC 180
Db 121 TTAGCTCCCAAGTTGGCATAGGAGGCCACAGTGGCTGAAATCATGCACTATTCGCTCGC 180
QY 181 GAAATTCCTGACTCCCGCGTAAACCCAAACCGTTCGAGGACAGGTTTTCCTGGATGACGGT 240
Db 181 GAAATTCCTGACTCCCGCGTAAACCCAAACCGTTCGAGGACAGGTTTTCCTGGATGACGGT 240
QY 241 TCCACAGGTTGCGAGGTGTTTCATCCGGCGCATCCACCGCGCTCCACGAGGCTCATGAG 300
Db 241 TCCACAGGTTGCGAGGTGTTTCATCCGGCGCATCCACCGCGCTCCACGAGGCTCATGAG 300
QY 301 CTGCGTCACGGTGGCGATCGCTACCTGGCAAGGGCGTTTGAAGGCGAGTTGAAAAAGTC 360
Db 301 CTGCGTCACGGTGGCGATCGCTACCTGGCAAGGGCGTTTGAAGGCGAGTTGAAAAAGTC 360
QY 361 AACGAGAAATCGCGCAGAGCTCGCTGGCTAGAGCTGACGATCAGATCAGCGCTCATCGAC 420
Db 361 AACGAGAAATCGCGCAGAGCTCGCTGGCTAGAGCTGACGATCAGATCAGCGCTCATCGAC 420
QY 421 GAACCAATGATCAAGCTTGTATGGCAGCGCAACAAAGTCCCGCTGGGTGCAAAACGCAATC 480
Db 421 GAACCAATGATCAAGCTTGTATGGCAGCGCAACAAAGTCCCGCTGGGTGCAAAACGCAATC 480
QY 481 CTTGGTGTTCCTATGGCTGTGTCGAAAGCTGCTGCTGCTGATTCGCGAGGCTCCCACTGTTT 540
```

Db 481 CTTGGTGTTCATGGCTGTTGCAAGGCTGCTGCTGATTCGCGAGGCTCCCACTGTTC 540
Qy 541 CGCTACATCGGTGACCAACGACACACGTTCTTCCAGTTTCCAATGATGAACATCATCAAC 600
Db 541 CGCTACATCGGTGACCAACGACACACGTTCTTCCAGTTTCCAATGATGAACATCATCAAC 600
Qy 601 GGTGGCGCTACGCTGACTCCGGTGTGACGTTTCCAGGAATTCATGATCGCTCCAAATCGGT 660
Db 601 GGTGGCGCTACGCTGACTCCGGTGTGACGTTTCCAGGAATTCATGATCGCTCCAAATCGGT 660
Qy 661 GCAGAGACCTTCTCTGAGGCTCTCGGCAACGCGGGGAGGTCTACGACGACTGAAGTCC 720
Db 661 GCAGAGACCTTCTCTGAGGCTCTCGGCAACGCGGGGAGGTCTACGACGACTGAAGTCC 720
Qy 721 GTCAATCAAGAAAAGGGCTCTGCAACGAGACTTGGCGATGAGGCGGCTTCGCTCTTCC 780
Db 721 GTCAATCAAGAAAAGGGCTCTGCAACGAGACTTGGCGATGAGGCGGCTTCGCTCTTCC 780
Qy 781 GTCGGCTCCACCGGTGAGGCTCTTGAACCTTATCGTTGAGGCAATCGAAGAGGCTGGCTTC 840
Db 781 GTCGGCTCCACCGGTGAGGCTCTTGAACCTTATCGTTGAGGCAATCGAAGAGGCTGGCTTC 840
Qy 841 ACCCGAGCAAGGACATCGCTCTTGTCTGAGAGTTCCTCTGAGTTCTTCAAGGAC 900
Db 841 ACCCGAGCAAGGACATCGCTCTTGTCTGAGAGTTCCTCTGAGTTCTTCAAGGAC 900
Qy 901 GGACACTTACCACTTCAAGGTTGGCGACGACCTCCGAGCTGAGATGGCAAGGTTTACGCT 960
Db 901 GGACACTTACCACTTCAAGGTTGGCGACGACCTCCGAGCTGAGATGGCAAGGTTTACGCT 960
Qy 961 GAGCTGTTGACGGTACCCCAATCGTCTCCATCGAGGACCCACCTGAGGAGTGCCTGG 1020
Db 961 GAGCTGTTGACGGTACCCCAATCGTCTCCATCGAGGACCCACCTGAGGAGTGCCTGG 1020
Qy 1021 GAGGTTTACCAACTTACCGCAACCATCGGCGACAAGGTTTCAAGTTCGTTGGCGAC 1080
Db 1021 GAGGTTTACCAACTTACCGCAACCATCGGCGACAAGGTTTCAAGTTCGTTGGCGAC 1080
Qy 1081 TTCTTGTGTCACCAACCTGAGGCTGGAAGGAGGCGATCGCTAAGAGGCTGCCAATCC 1140
Db 1081 TTCTTGTGTCACCAACCTGAGGCTGGAAGGAGGCGATCGCTAAGAGGCTGCCAATCC 1140
Qy 1141 ATCTGTGTTAAGGTGAACAGATCGGTACCTTCCACGAGACTTTCGACGCTGTGCAATG 1200
Db 1141 ATCTGTGTTAAGGTGAACAGATCGGTACCTTCCACGAGACTTTCGACGCTGTGCAATG 1200
Qy 1201 GCTCACCGCGAGGCTACACCTCCATGATGTCACCGTTCCCGTTCGAGACCGAGGACACC 1260
Db 1201 GCTCACCGCGAGGCTACACCTCCATGATGTCACCGTTCCCGTTCGAGACCGAGGACACC 1260
Qy 1261 ACCATTGCTGACCTCGCAGTTGCACTCACTGTGGCCAGATCAAGACTGGTGTCCAGCA 1320
Db 1261 ACCATTGCTGACCTCGCAGTTGCACTCACTGTGGCCAGATCAAGACTGGTGTCCAGCA 1320
Qy 1321 CGTTCCGACCGGTGCGCAAGTACAACAGGCTTCTCCGATCGAGGAGCTGCTTGGCGAC 1380
Db 1321 CGTTCCGACCGGTGCGCAAGTACAACAGGCTTCTCCGATCGAGGAGCTGCTTGGCGAC 1380
Qy 1381 GCCGGGCTTACGAGGTCGAGCGCATTTCCACGCTTTTCAGGGCTAATAAAGCGCTT 1440
Db 1381 GCCGGGCTTACGAGGTCGAGCGCATTTCCACGCTTTTCAGGGCTAATAAAGCGCTT 1440
Qy 1441 TTCCAGCCCGGTAACTCAAGTTGCGGGCGCTGTTGGCTTACTACTGTACTGTGT 1500
Db 1441 TTCCAGCCCGGTAACTCAAGTTGCGGGCGCTGTTGGCTTACTACTGTACTGTGT 1500
Qy 1501 GACTATGATCAGGATTTATGCAAGAGCAAGAAACCTCATAAAGGCTTCTTCTGTCT 1560
Db 1501 GACTATGATCAGGATTTATGCAAGAGCAAGAAACCTCATAAAGGCTTCTTCTGTCT 1560
Qy 1561 CAAGCAGGGAACGTGCTT 1578
|||||

Db 1561 CAAGCAGGGAACGTGCTT 1578
RESULT 3
US-09-103-840A-2
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

Query Match 40.1%; Score 632.6; DB 3; Length 4403765;
Best Local Similarity 68.4%; Pred. No. 9.4e-166;
Matches 891; Conservative 0; Mismatches 409; Indels 3; Gaps 1;
Qy 140 AGGAGGCGACAGTGGCTGAAATCATGACGATTCGCTGCGGAATTCCTGACTCCCGCG 199
Db 1144576 AGGAGAACCCAGTCCCGATTATCGAGCAGGTTAGGGCCCGAGAGATCTCGATTCCCGCG 1144635
Qy 200 GTAACCCAAACGTCGAGCGAGAGGTTTCTCTGAGATGACGTTCCACGCTGTCGAGGTG 259
Db 1144636 GCAACCCGACGCTGGAGTGGAGTGGCGCTTATCGAGGGACATTCGCCCGGCCCGCG 1144695
Qy 260 TTCCATCCGCGCATCCACCGGCTCCACGAGGCTCATGAGCTCGGTGACCGGTGGCGATC 319
Db 1144696 TCGCTCGGCGCTCCGACCGGGGAGCAGGAGCGCTCGAGTTGCGGACGCGCGATC 1144755
Qy 320 GCTACCTGGCGAAGGGGCTTTGAGGCGAGTTGAAACGTTCAAGAAATTCGCGACG 379
Db 1144756 GCTACGCGGCAAGGCGTGCAAAAGCCGTGCGAGGCTGTTTGTATGATGAGATCGGCCCG 1144815
Qy 380 AGCTCGCTGGCTAGAGGCTGACGATCAGCGCTTCATCGACGAAGCAATGATCAAGCTTG 439
Db 1144816 CCGTCATCGACTCAACGCGACGACGACGATTTGGTCGACGCGCTGGTGACCTAG 1144875
Qy 440 ATGCAACCGCCAAACAAGTCCCGCTGGGTGAAACGCAATCTTTGGTGTTCATGGCTG 499
Db 1144876 ACGCAACCCCGACAAGTCCCGCTGGGCGCAACGCGATCTTGGGTGTCGCTCGCTG 1144935
Qy 500 TTGCAAAAGGCTGTGCTGATTCGCGAGGCTCCACACTGTTCCGCTACATCGGTGGACCAA 559
Db 1144936 TTGCAAAAGGCGGCGGATTCGCGGAGCTGCGCTGTTTCCGTTATGTCGCGGGGCGCAA 1144995
Qy 560 ACGCACACGTTCTTCCAGTTTCCATGATGAACATCATCAACGCTGGGCTCACGCTGACT 619
Db 1144996 ACGGCGACATCTCCCGGTACCGATGATGAACATCTCAACGCGGGCGACACGCCGATA 1145055
Qy 620 CCGGTGTGAGTTTCAAGAAATTCATGATCGCTCCAAATCGGTGCGAGACCTTCTCTGAGG 679
Db 1145056 CCGGTGTGACATTCAGAGAGTTTCAATGTTGGCGCAAATTGGCGCGCCAGCTTCGTCGAGG 1145115
Qy 680 CTCTCCGCAACGCGCGGAGGTCTACCAAGCATGAAGTCCGTTCATCAAGGAAAGGGCC 739
Db 1145116 CGTTGCGCTGGGCGCTGAGGTGTACCAAGGCTCAAGTCGCTCAAGTCAAGGAGGCGG 1145175
|||||

740 TGTCCACCGGACTTGGCGATGAGGCGGCTTCGCTCCTTCGTCGCTCACCACCGTGAGG 799
 1145176 TGTCCACCGGCTGGGCGACGAGCGGCTTCGCCCGGATGTGGCGGACACCGCGG 1145235
 800 CTCTTGACCTTATCGTTGAGCAATCGAGAGGCTTGCTTACCCAGGCAAGACATCG 859
 1145236 CGTTGGACCTGATCAGCCGGGCGATCGAGTCGGCGGCTTCGAGACCGCGCGCGATGG 1145295
 860 CTCTTGCTCTGGAGCTTGTCTCTCTGAGTTCCTTCAAGGACGGAC---CTACACATTG 916
 1145296 CGTGGCCCTTGAGCGCGCGCGCCACCGAGTCTTTCACCGAGCGGACCGGCTACGTCCTG 1145355
 917 AAGTGGCCAGCACTCCGCGAGCTGAGATGGCAACGTTTACGCTGAGCTGTTGACCGCT 976
 1145356 AGGCAACACCCGTACCGCAGACGAGATGACCGAGTCTACCGCGGCTGCTCGCGGCT 1145415
 977 ACCCAATGCTTCATCGAGGACCCACTGCAAGGAGATGACTGGAGGGTTTACACCAACC 1036
 1145416 ACCGCTGGTGTGATCGAAGACCCACTGTCCGAAGACGATTTGGGACGGCTGGGCCGCG 1145475
 1037 TCACCGCAACCATCGGCGACAGGTTTTCAGATCGTTGGGAGGACTTCTTGTTCACCAACC 1096
 1145476 TGACGGCTCGATCGGTGACCGGGTGCAATCGTCGCGGACGACATCTTTGTCAACCAATC 1145535
 1097 CTGAGCGCTGAAGGAGGCGATCGCTAAGAAGGCTGCCAACTCCATCGTCTGTTAAGGTGA 1156
 1145536 CCGAGCGCTCGAGGAGGCGATCGAACGGGCGGTGGCAATGCTGCTGGTCAAGGTGA 1145595
 1157 ACCAGATCGGTACCCCTCACCGAGACCTTCGACGCTGTGACATGGCTCACCGGCGAGCT 1216
 1145596 ACCAGATCGGAGCGTTGACCGGAGACACTTCGACGCGGTCAAGCTGGCTCACCGCGCGAT 1145655
 1217 ACACCTCATGATGTCCTCCACCGTTCCGCTGAGACCGGAGGACACCACTTGTGACCTCG 1276
 1145656 ACCGACGATGATCAGTCAACCGAGTGGCGAGACGAGGACACCATGATCGCGACCTCG 1145715
 1277 CAGTTGCACTCAACTGTGGCGAGATCAAGACTGTGTCTCCAGACAGTTCCGACCGTGTG 1336
 1145716 CGGTGGCATCGGCGAGCGGCGAGATCAAGACGGGCGCGCTGCTCGCAGTGAGCGGTG 1145775
 1337 CAAGTACACCGAGCTTCTCGCATACGACGAGCTGTGTTGGGAGCGCGGCTTACGCGAG 1396
 1145776 CAATAACACAGCTGCTCGGATCGAAGAGGCGCTTGGCGAGCGCGCGCGCTACGCGG 1145835
 1397 GTCCGACGGATCCCGAGCTTTCAGGCTTAATAAAGCGCT 1439
 1145836 GCGACCTGGCATTTCTCCGTTTCGCTGCGAGACGAAATAGGT 1145878

RESULT 4
 US-09-103-840A-1
 ; Sequence 1, Application US/09103840A
 ; Patent No. 6294328
 ; GENERAL INFORMATION:
 ; APPLICANT: FLEISCHMAN, Robert D.
 ; APPLICANT: WHITE, Owen R.
 ; APPLICANT: FRASER, Claire M.
 ; APPLICANT: VENTER, John C.
 ; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
 ; TITLE OF INVENTION: TUBERCULOSIS
 ; FILE REFERENCE: 24366-20007.00
 ; CURRENT APPLICATION NUMBER: US/09/103,840A
 ; CURRENT FILING DATE: 1998-06-24
 ; NUMBER OF SEQ ID NOS: 2
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 1
 ; LENGTH: 4411529
 ; TYPE: DNA
 ; ORGANISM: Mycobacterium tuberculosis
 ; OTHER INFORMATION: H37Rv
 US-09-103-840A-1

Query Match 40.1%; Score 632.6; DB 3; Length 4411529;

Best Local Similarity 68.4%; Pred. No. 9 4e-166;
 Matches 891; Conservative 0; Mismatches 409; Indels 3; Gaps 1;
 140 AGAGGCGCAAGTGGCTGAAATCATGCACGTATTCGCTCGCGAAATTCATCGATCCCGCG 199
 1144551 AGGAGAAACCCAGTGCCGATTATCGAGCAGGTTAGGCGCCGAGAGATCCTCGATTCCCGCG 1144610
 200 GTAACCCAAACCGTCGAGGCGAGAGGTTTCTGTGATGACGGTCCCACGGTGTCCGAGGTG 259
 1144611 GCAACCGGACGGTGGAGTTCGAGGTGGCGCTTATCGACGGGACATTCGCCCGGCGCGG 1144670
 260 TTCCATCCGCGCATCCACCGGCTCCACGGGCTCATGAGCTCATGAGCTGCGTGCAGCGTGGCCATC 319
 1144671 TSCGTCGGCGGCTCGACCGGCGGACACGAGCCGCTCGAGTTGCGCGACGCGCGCGCATC 1144730
 320 GCTACCTTGGCGAAAGGGGTTTTGAAGGCACTTGAACAGTCAACAGAAAGTCAAGTGGGACG 379
 1144731 GCTACGCGCGCAAGGCGTGCAAAAAGCCGTGAGGCTGTTCTTGATGAGATCGGCCCGG 1144790
 380 AGCTCGCTGCGCTAGAGGCTGACGATCAGCGCTCATCGACGAGCAATCATCAAGCTTG 439
 1144791 CCGTCAATCGGACTCAACGCGCGACGACGCGATTTGGTCGACCGGCGCTGGTGACCTAG 1144850
 440 ATGCGCACCGCAACAGTCCCGCTGGGTGCAAAACGCAATCTTGGTGTTCATGCTGCTG 499
 1144851 ACGGCACCCCGACAGTCCCGCTGGCGGCAACGCGATCTTGGGTGTCTCGCTCGCTG 1144910
 500 TTGCAAGGCTGCTGTGATTTCCGAGGCTCCACATGTTCCGCTACATCGTGGGACAA 559
 1144911 TTGCCAAGGCGGCGGCGATTCGCGGAGCTGCGGTTTCCGTTATGTGGGGGGGCAA 1144970
 560 ACCGACACGTTCTTCAGATTCCCAATGATGAACATCATCAACGCTGGCGCTCAGCTCACT 619
 1144971 ACCGCGACATTCGCGGTACCGATGATGAACATCTCAACGCGGCGGCGACACCGCAT 1145030
 620 CCGGTGTGACGTTTCAAGGAAATCATGATCGCTCCAAATCGGTGCGAGACCTTCTCTGAGG 679
 1145031 CCGTGTGACATTTCAAGAGTTCATGTTGGCGCAATTTGGCGCGCCAGCTTCGTGAGG 1145090
 680 CTCTCGCAACCGGCGGAGGCTCTACACGACATGAAGTCCGTCATCAAGAAAGAGGCC 739
 1145091 CGTTGGCTGGGCGGCTGAGGTGTACCAACGCTCAAGTCGCTCTGAAAGAGAGGGG 1145150
 740 TGTCCACCGGACTTGGCGATGAGGCGGCTTCGCTCCTTCGCTCGGCTCCACCGCTGAGG 799
 1145151 TGTCCACCGGCTGGGCGAGGAGGCGCTTCGCCCGGATGTGCGCGGACCAACCGCGG 1145210
 800 CTCTTGACCTTATCGTTGAGGCAATCGAGAGGCTGGCTTTCACCCCGGCAAGGACATCG 859
 1145211 CGTTGACCTGATCAGCGCGGCGCATCGAGTGGCGGCTTCGACCGCGCGCGACGTG 1145270
 860 CTCTTGCTTGGAGCTTGTCTCTGAGTTCCTTCAAGGACGCGAC---CTACACCTTCG 916
 1145271 CGTGGCCCTTGGAGCGGCGGCGCCACGAGTTCCTTCAACCGGCGCGGCTACGCTCTCG 1145330
 917 AAGTGGCCAGCACTCCGCGAGCTGAGATGGCAACGTTTACGCTGAGCTGTTGACCGCT 976
 1145331 AGGCGACACCCGTACCGCAGACGAGTACCGAGTCTACGCGGCGCTGCTCGCGGCT 1145390
 977 ACCCAATGCTTCCATCGAGGACCCACTGCAAGGAGATGATGGAGGGTTTACACCAACC 1036
 1145391 ACCGCTGGTGTGATCGAAGACCCACTGTCCGAAGACGATTTGGAGCGGCTGGGCCCGCG 1145450
 1037 TCACCGCAACCATCGGCGACAGGTTTCAGATCGTTGGCGAGGACTTCTTGTCTACCAACC 1096
 1145451 TGACGGCTCGATCGGTGACCGGGTGCAATTCGTCGCGGACGACATCTTTGTCAACCAATC 1145510
 1097 CTGAGCGCTGAAGGAGGCGCATCGCTAAGAAGGCTGCCAACTCCATCCTGTTAAGGTGA 1156
 1145511 CCGAGCGGCTCGAGGAGGCGATCGAACGGGCGGTGGCAATGCTGCTGCTCAAGGTGA 1145570
 1157 ACCAGATCGGTACCCCTCACCGAGACCTTTCGACGCTGTGACATGGCTCACCGCGCGAGCT 1216

Db 1145571 ACCAGATCGGAGCGTTGACCGAGACATCGAGCGGCTCACGCTGCTCACCAACCGCGGAT 1145630
 Qy 1217 ACACCTCATGATGTCCACCGTTTCCGGTGAGACCGAGGACACCACTTCTGACCTCG 1276
 Db 1145631 ACCGACGATGATCAGTCAACCGAGTGGGAGAGACCACTGATCGCGACCTCG 1145690
 Qy 1277 CAGTTGCACTCACTGTGGCCAGATCAAGACTGGTGTCCAGACGTTTCCGACCGTGTG 1336
 Db 1145691 CGGTGGCCATCGGACGCGGAGATCAAGACGGCGCGCTGTGCGAGTGAGCGCGTGC 1145750
 Qy 1337 CAAAGTACAACACGCTTCTCCGCATCGAGCAGCTGTGGGAGCGCGCGCTGTACGCGAG 1396
 Db 1145751 CAAATACACACGCTGTGCGGATCGAAGAGCGCTTGGGAGCGCGCGCTGTACGCGG 1145810
 Qy 1397 GTCGACGCGCATCCACGCTTTCAGGCTTAAATAAAGCGCT 1439
 Db 1145811 GCGACCTGGCATTTCTCGGTTCGGTGTGCGAGACGAATAGGT 1145853

RESULT 5
 US-09-902-540-4780
 ; Sequence 4780, Application US/09902540
 ; Patent No. 6833447
 ; GENERAL INFORMATION:
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Wiegand, Roger C.
 ; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
 ; FILE REFERENCE: 38-10(15849)B
 ; CURRENT APPLICATION NUMBER: US/09/902,540
 ; CURRENT FILING DATE: 2001-07-10
 ; PRIOR APPLICATION NUMBER: 60/217,883
 ; PRIOR FILING DATE: 2000-07-10
 ; NUMBER OF SEQ ID NOS: 16825
 ; SEQ ID NO 4780
 ; LENGTH: 1299
 ; TYPE: DNA
 ; ORGANISM: Myxococcus xanthus
 US-09-902-540-4780

Query Match 35.0%; Score 552; DB 4; Length 1299;
 Best Local Similarity 65.7%; Pred. No. 6.3e-145;
 Matches 842; Conservative 0; Mismatches 425; Indels 15; Gaps 2;
 Qy 152 TGGCTGAAATCATGCAGTATTTCGTCGCGAAATTCGACTCCCGCGGTAAACCAACCG 211
 Db 2 TGACCGAGATTTCAGATTCTGGCGCGGAGTGTCTGACTCCCGTGGCAACCCGACCG 61
 Qy 212 TCGAGGCGAGGTTTCTCGATGACGGTTCCTCCAGGTGTGCGAGGTGTTCCATCCCGCG 271
 Db 62 TGGAGGCGGAGTCCAGCTTCGCGGGGCGCTCGTGGCGCGCGCGGTGCGGTGCGGTG 121
 Qy 272 CATCACCGGCGTCCAGAGGCTCATGAGTGTGCGTACGTTGGC---GATCGTACCTGG 328
 Db 122 CGTCCACCGGTGAGCATGAGGCCATCGAGTGTGCGCGACGCGGACAAAGCATCGCTACCTGG 181
 Qy 329 GCAAGGCGTTTGAAGCAGTTGAAACGTCACAGAGAAATCGGCGACGAGCTCGCTG 388
 Db 182 GCAAGGCGTGCAGAGGCGGTGAAGAACGTTGGTGCAGCTGCTGCGCGCGCGGTG 241
 Qy 389 GCCTAGAGGCTGAGCATCAGCGCTCATCGACGAGCAATGATCAAGCTTGTATGGCACCG 448
 Db 242 GCATGGAGCGCGGATCAGTTCCGGTGGACACGAGCGATGCTGAGCTGACCGGACCG 301
 Qy 449 CCAACAAAGTCCCGCTGGGTGCAACGCAATCTTGTGTTTCCATGCTGTTCGAAAG 508
 Db 302 CCACCAAGGCAAGTGGGCGCAACGCAATCTCGCGGTGTCCATGCGCGCGCTCGCG 361
 Qy 509 CTGCTGTGATTCGCGAGGCTTCCCATGTTCCGCTACATCGTGGGACCAACGACAG 568
 Db 362 CCGCGGCGGACGCGACGCGGTGCTGTTACCGCTAGCTGCGGCGCGGTGCGAGCGCGCA 421

Qy 569 TTCTTCAGTTTCAATGATGAACATCATCAACGCTGGCGCTCACGCTGACTCCCGTGTG 628
 Db 422 CCCTGCGGTCGCCCTGATGAACATCTCAACGCGCGCGCACCGGACACCGCGGTG 481
 Qy 629 ACCTTCAGGAATTCATGATCGCTCCAATCGGTGAGAGACCTTCTCTGAGGCTTCCGCA 688
 Db 482 ACCTGACGAGGTTTATGTTGTCGCCGCGGTGCTCTCTTCCGAGGCGCTTCCGCT 541
 Qy 689 ACAGCGCGGAGGTTTACACGCGCTGAAAGTCCGTCATCAAGGAAAGGCGCTTCCACCG 748
 Db 542 GGGGCGCGAGGTGTTCCACGCGCTGAAGAAGATTCTCAAGGCGCGCAAGCTGGCCACG 601
 Qy 749 GACTTGGCGATGAGGCGGCTTCCGCTCTTCCGCTGGCTCCACCGGTGAGGCTTGTACC 808
 Db 602 GCGTGGCGGACGAGGCGGCTTACGCCCGGACCTTCCGCGGAGCAAGAGGCGCTGAAGC 661
 Qy 809 TTATCGTTGAGGCAATCGAAGGCTGCTTCAACCGAGGCAAGGAGCATCGCTTGTGTC 868
 Db 662 TCATCATGGAGGCTCATCGAGCGGCTTCAAGCGGCTGAGGCTTCTTCTTGGGCC 721
 Qy 869 TGGAGCTTGTCTTCTGAGTTTCTTCAAGGACGCGCACTTACCACTT-----CG 916
 Db 722 TGGAGCTGGCGGCGAGGAGTTCTTTCGACAGGCGAGCAAGAGTACAAGCTGAAGGCG 781
 Qy 917 AAGTGGCCAGCACTCCGAGCTGAGATGGCAACGTTTACGCTGAGCTGTTGACGCT 976
 Db 782 AGGCAAGGAGTACGACTCGACCGGCTGCTCGAGTACTACCGGCGCTTCCGAGCGCT 841
 Qy 977 ACCCAATCGTCTCCATCGAGGACCACTGACGAGGAGTACTGGGAGGCTTACACCAAC 1036
 Db 842 ACCCATCATCTCCATCGAAGCGCATGCGGAGGATGACTGGGAGGCTTGAAGAAC 901
 Qy 1037 TCACCGCAACCATCGGCGCAAGGTTTCAGATCGTTGGCGAGCTTCTTCTGCTACCAAC 1096
 Db 902 TCACCGAGCGCTGGGTTTCGCGCATGAGTTGGTGGTGGAGCTTCTTCTGCTACCAAC 961
 Qy 1097 CTGAGCGCTGAAGGAGGCGCATCGCTAAGAGGCTGCCAATCCATCTGTTTAAAGTGA 1156
 Db 962 TGGAGCGCTGGGCGCGGCTCGAGACGCGGCAACTCCATCTCTGTTGAAGTGA 1021
 Qy 1157 ACCAGATCGGTACCTCACGAGACCTTCCAGCGTGTGACATGGCTCACCGCGAGCT 1216
 Db 1022 ACCAGATGGCACCTTACGAGAGCTTTCGACGCGTTCGATGGCGACCGCGCGCTT 1081
 Qy 1217 ACACCTCATGATGTCCACCGTTCGCTGAGACCGGAGGACCAACCATCTGCTGACCTCG 1276
 Db 1082 ACAGTCCGTCATGAGCAACGCTCCGCGAGACGAGGACCAACCATCTGCGGACCTGG 1141
 Qy 1277 CAGTTGCACTCACTGTGGCCAGATCAAGACTGGTGTCTCCAGACGTTTCCGACCGTGTG 1336
 Db 1142 CCGTGGCGTGGACTGCGGCGAGATCAAGACGGGTTCCGCGTCCCGCTCCGACCGCGTGG 1201
 Qy 1337 CAAAGTACAACGAGTCTTCGCGATCGAGAGCTGCTTGGGAGCGCGGCGCTTACGAG 1396
 Db 1202 CCAAGTACAACGAGTGTCTGCGATCGAAGCGAGCTGGGCGCGCGCGCTTACGCTG 1261
 Qy 1397 GTGCGAGCGCATCCCAAGCTT 1418
 Db 1262 GCAAGTGGTTCGCGCGCTT 1283

RESULT 6
 US-09-902-540-1235/c
 ; Sequence 1235, Application US/09902540
 ; Patent No. 6833447
 ; GENERAL INFORMATION:
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Wiegand, Roger C.
 ; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
 ; FILE REFERENCE: 38-10(15849)B
 ; CURRENT APPLICATION NUMBER: US/09/902,540

; CURRENT FILING DATE: 2001-07-10
 ; PRIOR APPLICATION NUMBER: 60/217,883
 ; PRIOR FILING DATE: 2000-07-10
 ; NUMBER OF SEQ ID NOS: 16825
 ; SEQ ID NO 1235
 ; LENGTH: 27903
 ; TYPE: DNA
 ; ORGANISM: Myxococcus xanthus
 US-09-902-540-1235

Query Match 35.0%; Score 552; DB 4; Length 27903;
 Best Local Similarity 65.7%; Pred. No. 3.1e-144;
 Matches 842; Conservative 0; Mismatches 425; Indels 15; Gaps 2;
 152 TGGCTGAATCATCGACGATTTCCTGATGACGGTTCCACGGTGTCCAGGTTTCATCCCGCG 211
 Db 1948 TGACCGAGATTTCAGATTTCGCGCGCGAAGTGTCTGACTCCCGGTTAAACCCACCG 19889
 Qy 212 TCAGGCGAGAGTTTCTCTGATGACGGTTCCACGGTGTCCAGGTTTCATCCCGCG 271
 Db 19888 TGAAGCGGAGGTCAGGTTGCGGGGGCGCTCGTGCCCGCGCGCGTCCGTTCCGTTG 19829
 Qy 272 CATCCACGGCGTCCACGAGGCTCATGAGTGTGCGTACGTTGC---GATGCTACCTGG 328
 Db 19828 CGTCCACGGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 19769
 Qy 329 GCAAGGGCGTTTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 388
 Db 19768 GCAAGGGCGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 19709
 Qy 389 GCCTAGAGGCTGACGATCAGCGCTCATCGACGAGCAATGATCAAGCTTGTATGGCACCG 448
 Db 19708 GCATGGACGCGCGGATCATGTTCCGGTGGACGAGCGAGTGTGAGCTGACGCGCACCG 19649
 Qy 449 CCAACAGTCCCGCTCGGTGCAACGCAATCTTGTGTTTCCATGCTGTGTTCAAGG 508
 Db 19648 CCACAAAGGCAAGCTGGCGGCAACGCAATCTTGTGTTTCCATGCTGTGTTCAAGG 19589
 Qy 509 CTGCTGTGATTCGCGAGGCTCCCATGTTCCGCTACATCGGTGACCAACGACGACG 568
 Db 19588 CCGCGGAGCGGACGCGGCTGCGCTGTACCGCTACGTTGGGGGGGTGCGAGCGGCA 19529
 Qy 569 TTCTTCCAGTTCCAAATGATGAACATCATCAACGGTGGCGTCACTGCTTCCGTTGTTG 628
 Db 19528 CCCTGCGGTGCGCTGATGAACATCTCAACGGCGGCGGACGCGGACGCGGCTGG 19469
 Qy 629 ACGTTACGAAATTCATGATCGCTCCATCGGTGAGAGACCTTCTCTGAGGCTTCCGCA 688
 Db 19468 ACGTGCAGGAGTTTCATGTTGTTGCGCGCGGCTGCTCTCTTCCGCGAGGCGCTGCGCT 19409
 Qy 689 ACGCGCGGAGGTTTACACGACTGAAGTCCGTCATCAAGGAAAGGGCTGTCCACCG 748
 Db 19408 GGGCGCGGAGGTTTCAACGCTGAGAGATCTCAAGGGCGGCAAGTGGCCACCG 19349
 Qy 749 GACTTGGCGATGAGGGCGGCTTCTCTTCCGTTCCGCTCCACCGTGAAGCTTCTTGAAC 808
 Db 19348 GCGTGGCGGAGGAGGCGGCTTACCGCGGAGCTTCCGCGGAGCAAGAGGCGCTGAAGC 19289
 Qy 809 TTATCGTTGAGGCAATCGAGAGGCTGCTTCCACCGGAGGAGGAGGAGGAGGAGGAGG 868
 Db 19288 TCATCATGGAGGCGATCGACGAGGCGGCTTCAAGGCGGAGGAGGAGGAGGAGGAGG 19229
 Qy 869 TGAAGTTGCTTCTCTGAGTTCTTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 916
 Db 19228 TGAAGTTGCTTCTCTGAGTTCTTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 19169
 Qy 917 AAGTGGCGGAGCTTCCGAGGCTGAGATGGAACAGTTTACGTTAGCTGAGCTGTTGACCGGT 976
 Db 19168 AGGCAAGGAGGAGTACGATCGACCGGCTGCTCGAGTACTACCGGGGCTCTCCGAGGCGCT 19109
 Qy 977 ACCCAATGCTTCCATCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1036
 Db 19108 ACCCAATGCTTCCATCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 19049

RESULT 7

US-09-252-991A-4449
 ; Sequence 4449, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; PRIOR FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 4449
 ; LENGTH: 1335
 ; TYPE: DNA
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-4449

Query Match 33.8%; Score 533.6; DB 4; Length 1335;
 Best Local Similarity 64.8%; Pred. No. 9.5e-140;
 Matches 848; Conservative 0; Mismatches 439; Indels 21; Gaps 3;
 131 AGTTGGCATAGGAGGCGACAGTGGCTGAATCATGACATGCTTCTCGTCCGGAATTTCTCG 190
 Db 26 ACTTCGAGGTGTTAAACAAGATGGCAAGATCGTCGACATCAAGGGCGGCTGAGTCTCGG 85
 Qy 191 ACTCCCGCGGTAAACCAACCGTCGAGCGAGAGGTTTTCTTGGATGACGGTTCCACCGGTG 250
 Db 86 ACTCCCGCGGCAACCTTACCGTTGAAGCGGACGCTGATCTCTGGACACGGCATCGTCGCA 145
 Qy 251 TCGCAGGTGTTTCATCCGCGCATTCACCGCGCTCCACGAGGCTCATGAGCTGGTGAACG 310
 Db 146 GCGCTCTCGCGGCTTCCGTTGCTTCCACCGGTTCCCGGAGGCGCTCGAGCTCGCGATG 205
 Qy 311 GTGGCGA---TCGCTACCTGGGCAAGGGCGTTTTTGAAGGAGGAGTTGAAAGCGTCAACCAAG 367
 Db 206 GCACCAAGAGCGGTTACCTGGGCAAGGGCGTGTGAAGCGGTGGCCCAACATCAACGCGC 265
 Qy 368 AAATCGGCGAGGAGCTCGCTGGGCTTAGAGGCTGACGATACGCGCTCATCGAAGAACAA 427

Db 1271 TGAGCGGTTCTGACCGGCTTCTAAATACAAACCAAGCTGATTTCGTATCGAAGAGCGCTGG 1330
Qy 1376 GCGAGCGCGGCGTCTACGAGCTCGCAGCGCATTCGCCAG 1415
Db 1331 GTGAGCAAGCCCGTTCAACGGTCTGAAGAGATCAAGG 1370

RESULT 10

US-09-818-780-22
; Sequence 22, Application US/09818780
; Patent No. 6677146
; GENERAL INFORMATION:
; APPLICANT: McHenry, Charles
; TITLE OF INVENTION: NOVEL THERMOPHILIC POLYMERASE III HOLOENZYME
; FILE REFERENCE: 1794.0030004
; CURRENT APPLICATION NUMBER: US/09/818,780
; CURRENT FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: US 60/192,736
; PRIOR FILING DATE: 2000-03-28
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 22
; LENGTH: 2363
; TYPE: DNA
; ORGANISM: Thermus thermophilus
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Includes 5' UTR and 3' UTR
US-09-818-780-22

Query Match 28.5%; Score 449.2; DB 4; Length 2363;
Best Local Similarity 64.3%; Pred. No. 7,1e-116;
Matches 692; Conservative 0; Mismatches 378; Indels 6; Gaps 1;

Qy 160 ATCATGACGATTCGCTCGCGAAATTCCTGACTCCCGCGGTAAACCCACCGTCGAGCA 219
Db 1279 ATCTCGCGGTCGCGGACGCGAGGTTTGGATTCCAGGGGCTTTCCACCGGTAGAGCG 1338
Qy 220 GAGGTTTCTCGGATGAGGTTCCACGGTCTCCAGGTGTTCCATCGCGGCGATCCACC 279
Db 1339 GAGGTGAGCTGGGAAGGGGGGCCAGGGGCCGCGGCGCATGGTGCCCTCCGGGGCTCCACC 1398
Qy 280 GCGCTCCACGAGGCTCATGAGCTGCTGACGGTGCGGATCGCTACCTGGGCAAGGGCGTT 339
Db 1399 GGAACCCAGAGGCGCTGAGCTCAGGACCGCGCGCAAGCGCTACCTGGGCAAGGGGGTG 1458
Qy 340 TTGAAGCAGTTGAAAACGTCAACGAAGAAATCGGCGACGAGCTCGCTGGGCTTAGAGCT 399
Db 1459 CGCGGGCGGTGGAGAACGTCAACGAGCGCATCGCCCCGAGCTCGTCGGCATGGAGCC 1518
Qy 400 GACGATGAGCGCTCATGACGAGCAATGATCAAGCTTGATGGCACCGCCAAAGTCC 459
Db 1519 CTGACACGAGGAGGGGTGAGCGGGCCATGCTGAGCTGGAGCGGACCGCCCAACAGGCC 1578
Qy 460 CGCTGGGTCAACGCAATCTTGGTGTTCATGGCTGTTGCAAGGGCTGCTGCTGAT 519
Db 1579 AACCTGGAGGAAGCGGCTCTCGCGTCTCCCTGGCCGTGGCCCGGCGGCGCGAG 1638
Qy 520 TCCGAGCGCTCCCACTGTTCCCGCTACATCGGTGGACCAACGACACAGTTCCTCCAGTT 579
Db 1639 GCGCTGGGCTGCCCCCTTTACCGCTACCTGGGCGGGTCCAGGGGGTCAACCTGCCCCGTG 1698
Qy 580 CCAATGATGACATCATCAACGGTGGGCGCTCACGCTGACTCCGGTGTGAGTTGAGGAA 639
Db 1699 CCCCTCATGACGTATCAACGGGGGGAAGCACCGCACACCGGGTGGACTTCCAGGAG 1758
Qy 640 TTATATCGCTCCAAATCGGTGACAGACCTTCTCTGAGGCTCTCCGCAACGGCGCGAG 699
Db 1759 TTATATGCTGGTCCCGGGGGCGGGAAGCTTCGCGAGGCTTGAGATCGGGCGCGAG 1818
Qy 700 GTCTACCAACGACGTGAAGTCCGTCATCAAGGAAAGGCGCTGTCCACGGGACTTGGCGAT 759

Db 1819 GTCTTCACACCCCTCAAGGGCGTCTCAAGAGAGGGGTACAGCACCAACGTCGGGGAC 1878
Qy 760 GAGGCGGCTTCGCTCTCTCCGTCGGCTCCACCGTCGAGGCTCTTGACCTTATCGTTGAG 819
Db 1879 GAGGAGGCTTCGCCCCCGACCTCAGGAGCAACAGAGGGCGGTGGAGCTTTTGTCTCCTC 1938
Qy 820 GCAATCGAAGAGGCTGCTTCAACCCAGGCAAGACATCGCTCTTGTCTCTGAGCGTTGCT 879
Db 1939 GCCATTGAGCGGGCGGGGTACACCCCGGCGCAGGAGGTCTCCCTGGCCCTGGACCCGGCC 1998
Qy 880 TCCTCTGAGTCTTCAAGGACGCGACCTACCACTTCGAAGTG-----GCCAGCACTCC 933
Db 1999 ACAGAGCGGCTTTACCCGGGACGGGAAGTACACCTGGAAGGGGAGGCAAGGTCTCTCTCC 2058
Qy 934 GCAGCTGAGATGGCAACGTTTACGCTCGCTGAGCTCGCTGACGGGTACCCAATCGTCTCCATC 993
Db 2059 TCGGAGGAGATGGTGGGCTTCTGGGAGCGCTGGGTGGAGAGTACCCCATCCGCTCCATT 2118
Qy 994 GAGGACCCACTGCGAGGAAGATGACTGGGAGGGTTACACCAACCTCACCGCAACCATCGGC 1053
Db 2119 GAGGAGCGCTTTCGCGAGGACGACTGGGAGGGGTGGCGGCTTCTCACCGAGCGCTGGGG 2178
Qy 1054 GACAAGTTTCAAGTCTGTTGGGAGGACTTCTTCGTACCAACCTCGAGCGCTGAAGGAG 1113
Db 2179 GGAAGGTTCAGCTCGTGGGGGACGACTCTTCTGTCACCAACCCGGAAGGCTCCGGGCG 2238
Qy 1114 GGCATCGCTAAGAGGCTGCCAATCCATCTGTTAAGGTGAACAGATCGGTACCCCTC 1173
Db 2239 GGGATTGAGCGGGGGTGGCCACGCCATCTTGTCAAGGTGAACAGATCGGAGCCCTC 2298
Qy 1174 ACCGAGACCTTCGACGCTGTGACATGGCTCACCGCGCAGGCTACACCTCCATGAT 1229
Db 2299 TCGGAAACCTTCGAGGCACTCCGCTGGCCAGGCTCGGGGTACAGGGCGGTGAT 2354

RESULT 11

US-08-781-802-9
; Sequence 9, Application US/08781802
; Patent No. 5969121
; GENERAL INFORMATION:
; APPLICANT: ALLEN, Larry
; APPLICANT: AIKENS, John
; APPLICANT: FONSTEIN, Michael
; APPLICANT: VONSTEIN, Veronika
; APPLICANT: DEMIRJIAN, David
; APPLICANT: CASADABAN, Malcolm
; TITLE OF INVENTION: Stable Biocatalysts for Ester Hydrolysis
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff
; STREET: 300 S. Wacker Drive 32nd Floor
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781,802
; FILING DATE: 10-JAN-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/694,078
; FILING DATE: 07-AUG-1996
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/019,580
; FILING DATE: 12-JUN-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/009,704

```

; FILING DATE: 11-JAN-1996
; PRIOR APPLICATION DATA: US 60/001,995
; FILING DATE: 01-AUG-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Chao, Mark
; REGISTRATION NUMBER: 37,293
; REFERENCE/DOCKET NUMBER: 95,963-E
; TELEPHONE: 312-913-0001
; TELEFAX: 312-913-0002
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3545 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: not relevant
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1397..2905
; OTHER INFORMATION: /note= "E019 sequence of longest
; OTHER INFORMATION: open reading frame; upstream untranslated region not exact"
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 1397..2905
; US-08-781-802-9

```

```

Query Match      28.0%; Score 442.4; DB 2; Length 3545;
Best Local Similarity 62.8%; Pred. No. 7.2e-114; Indels 15; Gaps 3;
Matches 743; Conservative 0; Mismatches 426;

Qy 229 CTGATGACGGTCCACGGTGTGCGCAGGTGTTCCATCCGGCGCATCCACGGCGTCCAC 288
Db 39 CGGGAAGAAGCGGTTTCGCCCGTGTAGTTCGCAAGCGCGTTTCGACGGCGAATAT 98

Qy 289 GAGCTCATGAGCTGCGTACGGTGGC---GATCGCTACCTGGCGAAGGGGTTTGAAG 345
Db 99 GAAGCGGTGAATTTGCGTGACGGCGACAAAACCGCTACCTCGGCAAAAGGGGTGCTCAA 158

Qy 346 GCAGTTGAAACGTCACAGGAATCGGCGACGAGCTCGCTGCGCTAGAGGCTGACGAT 405
Db 159 GCGGTTGAGAACGTCACGAAGTGTGCTCCGGAATCATCGGCTTAGAAGTACGTGAT 218

Qy 406 CAGCGCTCATCGACGAAGCAATGATCAAGCTTGATGGCACCGCAACAAAGTCCCGCTG 465
Db 219 CAAGTGGCGATCGACCGCGCTTGAATGAACCTTGACGGCACGGAAACAAAGGAAAGCTT 278

Qy 466 GGTGCAACGCAATCCTTGGTGTTCATGGCTGTTGCAAGGCTGCTGCTGATTCGCA 525
Db 279 GGGCGGAATGTAATTTAGGCGTGTGCTGCGCTGCGCTGCGCTGCGGCTGATGAGCTT 338

Qy 526 GGCCTCCACCTGTTCCGCTACATCGGTGGACCAAGCACACGTTCTTCCAGTTCCAATG 585
Db 339 GCGTTGCGGTGTACCAATACTTGGCGGCTTTTAAACGCTTAAACGCTGCTGACCGATG 398

Qy 586 ATGAACATCATCAACGGTGGCGCTCAGCTGACTCCGGTGTGTGACGTTTCAGGAATTCATG 645
Db 399 ATGAACATTTTAAACGGCGGCGCATGCGGACAAACAGCTTGTGACATTTCAAGAATTCATG 458

Qy 646 ATCGCTCCATCGTGCAGAGACCTTCTCTAGAGCTCTCCGCAACGGCGGAGGTTCTAC 705
Db 459 ATCATGCGGTCGTCGGGAAGCTTCGTTGAAGCGCTGCGCATGGGTGCGAGAATTTTC 518

Qy 706 CACGCACTGAAGTCCGTCATCAAGGAAAGGCGCTGTCACCGGACTTGGCGATGAGGCG 765
Db 519 CATAGCTTAAAGCTGTGTTAAAGCGAAGAGCTTACACACGCGTGTGCTGCGTACGAAGGC 578

Qy 766 GGCTTCGCTCTTCGTCGGGCTCCACCGTGGAGGCTCTTGACCTTATGTTGAGGCAATC 825
Db 579 GGATTTGCTCGAACTTAAATTCGAACGAAGAAGCGCTGCAAAACGATCATTTGAAGCGATC 638

```

RESULT 12

```

US-09-583-110-1156
; Sequence 1156, Application US/09583110
; Patent No. 6699703
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; TITLE OF INVENTION: Pneumoniae for Diagnostics and Therapeutics
; FILE REFERENCE: PATH00-07A
; CURRENT APPLICATION NUMBER: US/09/583,110
; PRIOR FILING DATE: 2000-05-26
; CURRENT APPLICATION NUMBER: US 09/107,433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 1156
; LENGTH: 1305
; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae
; US-09-583-110-1156

```

```

Query Match      27.5%; Score 433.2; DB 4; Length 1305;
Best Local Similarity 59.9%; Pred. No. 1.7e-111;
Matches 776; Conservative 0; Mismatches 493; Indels 27; Gaps 2;

Qy 160 ATATGACGCTATTCGCTCGCGAAATTCCTGACTCCCGCGTAACCCAAACGTCGAGCA 219
Db 10 ATTACTGATGTTTACGCTCGGAAGTCTTAGACTCAGCGGTAAACCAACTTGAAGTA 69

```



```

Db 186 TTAGCGGAACCTCTTAACAAAGGTAATATTAGTGGCAACGCTATTTTAGTGGTTTCTATC 245
Qy 496 GCTGTTGCAAAAGGCTGTCTGATGTTCCGCGAGGCTCCCACTGTTCCGCTACATCGGTGGA 555
Db 246 GCTGTTGCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 305
Qy 556 CCAAAACGCAACGCTTCTTCGATGTTCCAAATGATGAACATCATCAACGCGGATCATCGCT 615
Db 306 TTCAACACAAAGATTTGCAACCAACCAATGATGAACATCATCAACGCGGATCATCGCT 365
Qy 616 GACTCCGGTGTGACGTTGACGTTGACGTTGACGTTGACGTTGACGTTGACGTTGACGTTG 675
Db 366 GATAACTCAATCGACTTCCAAAGATTCATGATCATGCTGTTGGCGTCCAACTTCAAA 425
Qy 676 GAGGCTCTCCGCAACGCGCGGAGTCTTACCAACGCTGAAAGTCCGCTCATCAAGGAAAG 735
Db 426 GAAGCTCTACGATGCGTGAAGTATTCACGCAATGCTTCAATCTTGAAGCTCGC 485
Qy 736 GGCCTGTCCACCGGACTTTGGCGATGAGGGCGCTTTCGCTCCTTCGTCGGCTCCACCGT 795
Db 486 GGATTAGCAACTTCTGTTGGTGACGAAAGGTGCAATTTGCTCCAAACCTTGGTTCAACGAA 545
Qy 796 GAGGCTCTGACCTTATCGTTGAGCAATCGAAGGCTGGCTTACCCAGGCAAGGAC 855
Db 546 GAAGGTTTGAAGTAAATCATCGAAGCAATCGAAGGCTGGCTATGTACCTGGTAAAGAC 605
Qy 856 ATCGCTCTTGTCTGGAGCTTGTCTCTCTGAGTTCTTCAAGGAGCGGCACTACCACTTC 915
Db 606 GTTGTCTTGTATGGATGCTGCTTCTTCAGAAATCTACGCAAGAAAGGTTGTTAC 665
Qy 916 GAA-----GTTGGCGAGCACTCCGCGAGTGAGTGGAAGGTTTACGCT 960
Db 666 GTACTAGCTGATTGAGGCGAAGGCGAAGGCAAAACAACTGACGAAATGATCAAAATCTACGAA 725
Qy 961 GAGCTCTGACGCTACCAATCGTCTCCATCGAGGACCCACTGCGAGGATGACTGG 1020
Db 726 GAATTAGTTTCTAAATATCCAAATCATCTCAATCGAAGACGGATTAGACGAAACGACTGG 785
Qy 1021 GAGGTTTACACCAACCTCACCGCAACCAATCGGCGCAAGGTTTCAGATGTTGGCGAGCAG 1080
Db 786 GACGGATTCAAGAAATTAACGTGAGTATTAGCGCAAGAAAGTTCAATTAGTTGGTGAGCAG 845
Qy 1081 TTCTTGTGTCACCAACCTGAGCGCTGAGGAGGCGATCGCTAAGAAAGGCTGCCAACTCC 1140
Db 846 TTGTTGCTGTAACAAACACTCAAAATATATCTGAAGTATCGAAGAAAGGAAATGCTAACTCA 905
Qy 1141 ATCTGTTTAAAGTGAACAGATCGGTACCTCACGAGACCTTCGACGCTGTGCAATG 1200
Db 906 ATCTTATCAAGTGAACCAATCGGTACATTAACAGAAACTTTTGAAGCTATTCGAAATG 965
Qy 1201 GCTCACCGCGAGGCTACACCTCCATGATGTCACACGCTTCCGCTGAGACCGAGGACAC 1260
Db 966 GCTAAGAGCTGCTACAGAGGTTGATATCACCGTTCTGGTGAACAGAAAGATTCA 1025
Qy 1261 ACCATTGCTGACCTCGCATGTTGCACTCAACTGTGGCCAGATCAAGACTGGTGTCTCCAGCA 1320
Db 1026 ACAATCTCTGATATCGCTGTTGCAACAAACGCTGGCCAAATCAAAACTGTTCTCTATCA 1085
Qy 1321 GGTTCGACCGTGTGCAAGATACACAGGCTTCTCCGATCGAGAGCTGCTTGGCGAC 1380
Db 1086 CGTACTGACCGGTATGCTTAAATACAAACCAATTAATAGTATCGAAGACCAACTTGGTGAA 1145
Qy 1381 GCGGGGTCTACGAGGTGCGAGGCAATCCACGCTTTCAGGGCTAA 1428
Db 1146 GTTCAGAAATCAAAAGGTTTGAATCTTTTACAACTTAAACAAACAA 1193

```

RESULT 15

US-09-489-039A-2858/c
; Sequence 2858, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al

```

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 2858
; LENGTH: 1233
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
; US-09-489-039A-2858

```

```

Query Match 26.6%; Score 420.2; DB 4; Length 1233;
Best Local Similarity 62.0%; Pred. No. 7.3e-108;
Matches 728; Conservative 0; Mismatches 423; Indels 24; Gaps 3;

```

```

Qy 265 TCCGCGCATCCACCGCGTCCACGAGGCTCATGAGTGTGCGTGACGGTGGCGA---TCGC 321
Db 1233 TCAGGTGCTTCTACTGTTCCCGCAAGCGCTGGAACCTGCGGATGGCGCAAAATCCCGT 1174
Qy 322 TACCTGGCAAGGCGCTTTGAAGGCAAGTTGAAACCGTCAACGAGAAATCGGCGACGAG 381
Db 1173 TTCTGGGTAAAGCGCTAAACCAAGCTGTTGCTCGGTAAACGCGCCGATCGCTCAGGCA 1114
Qy 382 CTGCTCGGCTAGAGGCTGACGATCAGCGCTCATCGACGAAGCAATGATCAAGCTTCAT 441
Db 1113 ATCTGGGCAAGATGCTTAAGATCAGGCTGGCATCCAAAGATCATGATCGACTGGAC 1054
Qy 442 GGCACCGCCCAACAGTCCCCTGGGTGCAAAACGCAATCTCTTGGTGTGTTTCCATGGCTGT 501
Db 1053 GGTACTGAAACAAATCTAACTTCGGTGGCAACGCTATCTCTGGCGTTTTTCTTGCGCAAC 994
Qy 502 GCAAGGCTGCTGCTGATTCGGGAGGCTCCCACTGTTCCGCTACATCGGTGGACCAAC 561
Db 993 GCCAAGCAGCAGCTGCTGCCGTCGCGTCCAAAGGCTGCGCTGTACGCGCACATCGCTGAACTGAAC 934
Qy 562 GC-----ACAGTTCTTCCAGTTCCCAATGATGAACATCATCAACGCGTGGCGCT 609
Db 933 GGCACCTCGGCAAAATTAATCTCATGCGGTTCCGATGATGAACATCATCAACGCGGTGAG 874
Qy 610 CAGCTGACTCCGGTGTGAGCTTTCAGGAAATTCATGATCGCTCCAATTCGGTGCAGAGACC 669
Db 873 CAGCTGACAAACAGCTGACATCCAGGAAATTCATGATTCAGCGGTTGGCGCGCCGACC 814
Qy 670 TTCTCTAGGCTCTCCGCAACGCGCGAGGTCTACCAAGCACTGAAGTCTGATCAAG 729
Db 813 CTGAAAGAGCAGTACGCTGGGTTCTGAAAGTGTTCATCATCACTTGGCGAAAAGTGTGAAG 754
Qy 730 GAAAGGCGCTGTCACCGGACTTTGGCGATGAGGCGGCTTTCGCTCCTTCGCTCGGCTCC 789
Db 753 TCCAAAGCATGAACATCTGCGTTGGTGAAGGAGGCTACGCGCGAACTGGGTTC 694
Qy 790 ACCGCTGAGGCTCTTGACCTTATCGTTGAGGCAATCGAGAGGCTGCTTCCACCCAGGC 849
Db 693 AACGCCAAGCGCTGGCTGTTATCGCTGAAAGCGGTTAAAGCTGCAGGCTACGAGCTGGGC 634
Qy 850 AAGGACATCGCTCTTGTCTGAGCTGCTTCTCTGAGTCTTCAAGGAGCGGACCTAC 909
Db 633 AAAGACATCACTCCCTGGCGATGGAATGCGCGCGCTCTGAAATTTCTACAAAGACGCTAAATAC 574
Qy 910 CACTTCCGAAGTGGC-----CAGCACTCCGCGCTGAGATGGGCAACGTTTACGCT 960
Db 573 GTTCTGGCTGGCAAGGCAACAAAGCGTTTCACTCTTGAAGAGTTCACTCACTTCTCTGAA 514
Qy 961 GAGCTCGTTGACGCGTACCCAAATCGTCTCCATCGAGGACCCCACTGCGAGGAAGATGACTGG 1020
Db 513 GAGCTGACCAACAGTACCCGATGCTCTCCATCGAAGATGCTCTGGAGCAATCTGACTGG 454
Qy 1021 GAGGTTTACCAACCTTCAACGCAACCATCGGCGCAAGGTTTCAAGTCTGTTGGCGACGAC 1080
Db 453 GAAGGTTTTCGCTTACCAGACTTAAAGTACTGGGCGCAAAATCCAGCTGTTGGTGGTGGCAGC 394

```

Qy	1081	TTCTTCGTACCAACCTGAGCGCTGAGAGGGGATCGCTAAGAAGGCTGCCAACTCC	1140
Db	393	CTGTTCGTAAACCAACACCAAGATCCTGAAAGAAGGCATCGAAAAAGGCATCGCTAACTCC	334
Qy	1141	ATCCTGGTTAAGGTGAACCAAGATCGGTACCTCACCGAGACCTTCGACGCTGTGACATG	1200
Db	333	ATCCTGATCAAGTTCACCAAGATCGGTTCCTGACCGAAGCTTGGCTGCTATCAAGATG	274
Qy	1201	GCTCACCGGAGGCTACACCTCCATGATGTCCACCGTTCGGTGAGAGCCGAGGACCC	1260
Db	273	GCGAAAGACGTGGCTACACCGCTGTTATCTCTCACCGTTCGGCGAAACTGAAGACGCT	214
Qy	1261	ACGATTGCTGACCTCGCAGTTGCACCTCAACTGTGGCCAGATCAAGACTGGTGCTCCAGCA	1320
Db	213	ACGATTGCCGACCTGGCTGTGTGTACCGCTGACGCCGAGATCAAAACTGGTTCCTATGAGC	154
Qy	1321	CGTTCGACCGTGTGCGAAAGTACAACAGCTTCTCGGCATCGAGCAGCTTGGCGAC	1380
Db	153	CGTTCGACCGGTGTGCTAAATACAACAGCTGATTGATTCGAAGAAGCGCTGGTGAG	94
Qy	1381	GCCGGCTCTACGAGGTGCGAGCGCATTTCCACG	1415
Db	93	CAAGCGCGTTCAACGCTCGTAAGAGATCAAAGG	59

Search completed: September 28, 2005, 23:33:36
 Job time : 294.5 sec

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 28, 2005, 16:02:19 ; Search time 6707 Seconds
(without alignments)
11400.377 Million cell updates/sec

Title: 10:28:47 2005
Perfect score: 1578
Sequence: 1'ggctgggagataggtagtt.....ctcaagcagggaacgtgctt 1578

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*
1: gb_ba.*
2: gb_hg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1578	100.0	1578	6 AR490885	AR490885 Sequence
2	1578	100.0	1578	6 AX136862	AX136862 Sequence
3	1578	100.0	1578	6 BD014844	BD014844 Polynucle
4	1578	100.0	333150	1 AP005277	AP005277 Coryneb
5	1578	100.0	348475	1 BX927150	BX927150 Coryneb
6	1578	100.0	349980	6 AX127146	AX127146 Sequence
7	1576.4	99.9	1578	6 AR490886	AR490886 Sequence
8	1403.4	88.3	1405	6 AX763190	AX763190 Sequence
9	1396.4	88.5	1398	6 AX064945	AX064945 Sequence
10	1275	80.8	1275	6 BD163286	BD163286 Novel pol
11	1275	80.8	1275	6 AX121169	AX121169 Sequence
12	1036.4	65.7	300750	1 AP005217	AP005217 Coryneb
13	917.4	58.1	347625	1 BX248356	BX248356 Coryneb
14	734	46.5	276800	1 SC039115	SC039115 Streptomy
15	701.6	44.5	110000	1 AE016822	Continuation (18 o
16	698.4	44.3	296300	1 AP005035	AP005035 Streptomy
17	680.6	43.1	110000	1 AP006618	Continuation (51 o
18	657.6	41.7	303855	1 AE017230	AE017230 Mycobacte
19	642.4	40.7	30159	6 CQ363776	CQ363776 Sequence

20	642.4	40.7	110000	1 AE017283_05	Continuation (6 of
21	632.6	40.1	45191	6 AX704278	AX704278 Sequence
22	632.6	40.1	110000	1 AE000516_11	Continuation (12 o
23	632.6	40.1	327650	1 BX248337	BX248337 Mycobacte
24	632.6	40.1	349306	1 BX842575	BX842575 Mycobacte
25	588.4	37.3	329100	1 SMES91787	AL591787 Sinorhizo
26	584.6	37.0	10029	1 AE014725	AE014725 Bifidobac
27	584.6	37.0	349980	6 AX492784	AX492784 Sequence
28	584.6	37.0	349980	6 AX492785	AX492785 Sequence
29	584.6	37.0	349980	6 AX553951	AX553951 Sequence
30	584.6	37.0	349980	6 AX553952	AX553952 Sequence
31	583.4	37.0	344050	1 MLEPRTN1	AL583917 Mycobacte
32	577.8	36.6	12058	1 AE009103	AE009103 Agrobacte
33	577.8	36.6	12407	1 AE008068	AE008068 Agrobacte
34	577	36.6	347660	1 AP002994	AP002994 Mesorhizo
35	576.4	36.5	10934	1 AE005846	AE005846 Caulobact
36	576.4	36.5	12709	1 AE014414	AE014414 Brucella
37	573.2	36.3	13776	1 AE009526	AE009526 Brucella
38	571.8	36.2	1290	1 AY181039	AY181039 Methyloba
39	571.4	36.2	349981	1 BX572602	BX572602 Rhodopseu
40	558	35.4	300600	1 AP005952	AP005952 Bradyrhiz
41	557.8	35.3	193050	1 AL646062	AL646062 Ralstonia
42	554	35.1	197615	5 BX470161	BX470161 Zebrafish
43	551.8	35.0	110000	1 AE017180_24	Continuation (25 o
44	551.8	35.0	110000	1 AE017180_25	Continuation (26 o
45	543.6	34.4	110000	1 AP006840_02	Continuation (3 of

ALIGNMENTS

RESULT 1

AR490885

LOCUS

Sequence 1 from patent US 6713289.

AR490885

ACCESSION

AR490885

VERSION

AR490885.1

GI:47258404

KEYWORDS

Unknown.

SOURCE

Unknown.

REFERENCE

1 (bases 1 to 1578)

Mockel, B., Pfeifferle, W., Hermann, T., Puhler, A., Kalinowski, J. and

Bathe, B.

Patent: US 6,713,289 A, 30-MAR-2004;

Location/Qualifiers

1..1578

source

/organism="unknown"

/mol_type="genomic DNA"

ORIGIN

Query Match

Best Local Similarity 100.0%; Score 1578; DB 6; Length 1578;

Matches 1578; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GGCTGGGATATGGGTAGTATTTCCGACACTAATTTCACTGCTCATCGAACAAGA 60

1 GGCTGGGATATGGGTAGTATTTCCGCCACTAATTTCACTGCTCATCGAACAAGA 60

61 TTTCGTCACAAATTTGGGTAGACGATTTGAAGACATTTTGAATCACTGATTAATCTAG 120

61 TTTCGTCACAAATTTGGGTAGACGATTTGAAGACATTTTGAATCACTGATTAATCTAG 120

121 TTAGCTCTCCCAAGTTGGCATAGGAGGCCACAGTGGCTGAAATCATGACGATTCGCTCGC 180

121 TTAGCTCTCCCAAGTTGGCATAGGAGGCCACAGTGGCTGAAATCATGACGATTCGCTCGC 180

181 GAAATTTCTGACTCCCGCGGTAAACCAACCGTCGAGCAGAGGTTTCTGGATGACGGT 240

181 GAAATTTCTGACTCCCGCGGTAAACCAACCGTCGAGCAGAGGTTTCTGGATGACGGT 240

241 TCCCACGGTGTCCGAGGTGTTCCATCCGGCGCATCCACCGCGCTCCACGAGGCTCATGAG 300

```
Db 241 TCCACGGTCTCGAGGTTCCATCGGCGCATCCACCGCGCTCCACGAGGCTCATGAG 300
Qy 301 CTGCGTGACGGTGGCGATCGCTACCTGTGGCAAGGGCGTTTGAAGGCGATTGAAAACGTC 360
Db 301 CTGCGTGACGGTGGCGATCGCTACCTGTGGCAAGGGCGTTTGAAGGCGATTGAAAACGTC 360
Qy 361 AACGAGAAATCGGCGACGAGCTCGCTGGCTAGAGGCTGACGATCAGCGCCTCATCGAC 420
Db 361 AACGAGAAATCGGCGACGAGCTCGCTGGCTAGAGGCTGACGATCAGCGCCTCATCGAC 420
Qy 421 GAAGCAATGATCAAGCTTGTATGGCACCGCCCAACAAAGTCCCGCTCGGTGCAAAACGCAATC 480
Db 421 GAAGCAATGATCAAGCTTGTATGGCACCGCCCAACAAAGTCCCGCTCGGTGCAAAACGCAATC 480
Qy 481 CTTGGTGTTCATGGCTGTGTTGAAAGGCTGTGCTGATTCGGAGGCGCTCCCACTGTTTC 540
Db 481 CTTGGTGTTCATGGCTGTGTTGAAAGGCTGTGCTGATTCGGAGGCGCTCCCACTGTTTC 540
Qy 541 CGCTACATCGGTGGACCAACGACACGTTCTTCCAGTTCCAATGATGACATCATCAAC 600
Db 541 CGCTACATCGGTGGACCAACGACACGTTCTTCCAGTTCCAATGATGACATCATCAAC 600
Qy 601 GGTGGCGCTCACGCTGACTCGGTGTTGACGTTTACGAAATTCATGATCGCTCCAATCGGT 660
Db 601 GGTGGCGCTCACGCTGACTCGGTGTTGACGTTTACGAAATTCATGATCGCTCCAATCGGT 660
Qy 661 GCAGAGACCTTCTGTAGGCTTCCGCAACGGCGCGAGGTGTACACGCACTGAAGTCC 720
Db 661 GCAGAGACCTTCTGTAGGCTTCCGCAACGGCGCGAGGTGTACACGCACTGAAGTCC 720
Qy 721 GTCATCAAGGAAAGGCGCTTCCACCGACTTGGCGATGAGGCGGCTTCGCTCCTTCC 780
Db 721 GTCATCAAGGAAAGGCGCTTCCACCGACTTGGCGATGAGGCGGCTTCGCTCCTTCC 780
Qy 781 GTGGCTCCACCGGTGAGGCTCTTGACCTTATCGTTGAGGCAATTCGAAAGGCTTGGCTTC 840
Db 781 GTGGCTCCACCGGTGAGGCTCTTGACCTTATCGTTGAGGCAATTCGAAAGGCTTGGCTTC 840
Qy 841 ACCCAGGCAAGGACATCGCTCTTGTCTGTGAGCGTTGCTTCTCTGTGATTTCTCAAGGAC 900
Db 841 ACCCAGGCAAGGACATCGCTCTTGTCTGTGAGCGTTGCTTCTCTGTGATTTCTCAAGGAC 900
Qy 901 GGCACTTACCACCTCGAAGGTGGCCAGCACTCCGCACTGAGATGGCAACGTTTACGCT 960
Db 901 GGCACTTACCACCTCGAAGGTGGCCAGCACTCCGCACTGAGATGGCAACGTTTACGCT 960
Qy 961 GAGCTCGTTGACGGTACCCAAATCGTCTCCATCGAGGACCCACTGCGAGGAAGATGACTGG 1020
Db 961 GAGCTCGTTGACGGTACCCAAATCGTCTCCATCGAGGACCCACTGCGAGGAAGATGACTGG 1020
Qy 1021 GAGGTTTACACCAACCTCAACGCAACCATCGGCGACAAGGTTTCAGATCGTTGGCGACGAC 1080
Db 1021 GAGGTTTACACCAACCTCAACGCAACCATCGGCGACAAGGTTTCAGATCGTTGGCGACGAC 1080
Qy 1081 TTCTTCGTTCACCAACCTCGAGGCTGAGGCGCTGAGGCGGATCGTAAAGGCTGCGCACTCC 1140
Db 1081 TTCTTCGTTCACCAACCTCGAGGCTGAGGCGGATCGTAAAGGCTGCGCACTCC 1140
Qy 1141 ATCTCGTTAAGGTGAACCCAGATCGGTACCTTACCGAGACCTTCGACGCTGTGACATG 1200
Db 1141 ATCTCGTTAAGGTGAACCCAGATCGGTACCTTACCGAGACCTTCGACGCTGTGACATG 1200
Qy 1201 GCTCACCGCGAGGCTACACCTCCATGATGTCCACCGTTCGCGTGAAGCGGAGCAC 1260
Db 1201 GCTCACCGCGAGGCTACACCTCCATGATGTCCACCGTTCGCGTGAAGCGGAGCAC 1260
Qy 1261 ACCATTGCTGACCTCGAGTTGCACTCACTGTGGCGAGATCAGACTGGTGTCTCCAGCA 1320
Db 1261 ACCATTGCTGACCTCGAGTTGCACTCACTGTGGCGAGATCAGACTGGTGTCTCCAGCA 1320
Qy 1321 CGTTCCGACCGGTGCGCAAAAGTACAACGAGCTTCTCCGCGATTCGAGGAGCTGCTTGGCGAC 1380
```

```
Db 1321 CGTTCCGACCGGTGTCGCAAAAGTACAACCAAGCTTCTCCGATCCAGCAGCTGCTTGGCGAC 1380
Qy 1381 GCGGGGCTTACGCGAGGTCGACGCGCATTCGCCAGCTTTTCCAGGGCTAAATAAAGCGCTT 1440
Db 1381 GCGGGGCTTACGCGAGGTCGACGCGCATTCGCCAGCTTTTCCAGGGCTAAATAAAGCGCTT 1440
Qy 1441 TTTCGACGCCCGGTAAACCTCAAGGTTGCCGGGCGTGGTTGCCCTTACTACTGTTACTGGTGT 1500
Db 1441 TTTCGACGCCCGGTAAACCTCAAGGTTGCCGGGCGTGGTTGCCCTTACTACTGTTACTGGTGT 1500
Qy 1501 GACTATGATCGAGGATTTATGGCAAAACGCAAGAACTCATAAAGGCTTTGTTCTGTCT 1560
Db 1501 GACTATGATCGAGGATTTATGGCAAAACGCAAGAACTCATAAAGGCTTTGTTCTGTCT 1560
Qy 1561 CAAGCAGGGAACGTCGTT 1578
Db 1561 CAAGCAGGGAACGTCGTT 1578

RESULT 2
AX136862
LOCUS AX136862 1578 bp DNA linear PAT 30-MAY-2001
DEFINITION Sequence 1 from Patent EP1090998.
ACCESSION AX136862
VERSION AX136862.1 GI:14273213
KEYWORDS Corynebacterium glutamicum
SOURCE Corynebacterium glutamicum
ORGANISM Corynebacterium glutamicum
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Corynebacteriaceae; Corynebacterium.

REFERENCE 1
AUTHORS Moeckel,B., Pfeifferle,W., Hermann,T., Puehler,A., Kalinowski,J. and
Bathe,B.
TITLE Nucleotide sequences coding for the eno gene
JOURNAL Patent: EP 1090998-A 1 11-APR-2001;
Degussa AG (DE)
FEATURES
    Location/Qualifiers
        1..1578
            /organism="Corynebacterium glutamicum"
            /mol_type="unassigned DNA"
            /db_xref="taxon:1718"
            151..1428
                /note="unnamed protein product"
                /codon_start=1
                /transl_table=11
                /protein_id="CAC39860.1"
                /db_xref="GI:14273214"

CDS
    151..1428
        /translation="MAEIMHVFAREILDSRGNPTVEAEVFLDDGSHGVAGVPSGASTG
        VHEAHLRGGDRYLKGLVKAVENNERIGDELAGEADQRLIDAEIMIKLDGTANK
        SRLGANIIIGVSMAYAKAADSAGLPFRVIGGNPAHLVLPVPMNIIINGAHADSGVD
        VQEFMLAPIGAETTFSEALRNGAEVYHALKSVIKEGLSTGLDGGEGFAPSVGSTREAL
        DLIVEAIEKAGFTPGKDIALADVASSEFFKDGTHFEGGQHSAAEMANVYARLVDAV
        PIVSTEDPQEDDWBYNTIATIDGKQIVGDDFFVTNPERLKEGIIAKKAANSILVK
        VNQIGTLTTFDAVMHRAGYTSMMSHRSGETDTTIADLAVLNCGQIKTGAPARS
        DRVAKYNQLRLRIQLLDGAVGVAGRSAPFRFQG"

ORIGIN
Query Match 100.0%; Score 1578; DB 6; Length 1578;
Best Local Similarity 100.0%; Pred. No. 6.9e-288;
Matches 1578; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCGTGGGATATGGGTAGTTTTCGCCACTAAATTTCAACTGATTCCTCATCGAAAACAAG 60
Db 1 GCGTGGGATATGGGTAGTTTTCGCCACTAAATTTCAACTGATTCCTCATCGAAAACAAG 60
Qy 61 TTTCGTCAACAATTTGGGTGTAGACGTGATTGAAGACATTTGATCATCGTGAATAATTTCTAG 120
Db 61 TTTCGTCAACAATTTGGGTGTAGACGTGATTGAAGACATTTGATCATCGTGAATAATTTCTAG 120
Qy 121 TTAGTCTCCCAAGTTGGCATAGGAGGCCACAGTGGCTGAAATCATGCACGTATTCGCTCGC 180
Db 121 TTAGTCTCCCAAGTTGGCATAGGAGGCCACAGTGGCTGAAATCATGCACGTATTCGCTCGC 180
```

QY 181 GAAATTCCTGACTCCCGGGTAACCAACCGTTCAGGCGAGAGGTTTCTCGATGACGGT 240
 Db 181 GAAATTCCTGACTCCCGGGTAACCAACCGTTCAGGCGAGAGGTTTCTCGATGACGGT 240
 QY 241 TCCACAGGTTGCGCAGGTGTTCCATCCGCGCATCCACCGCGTCCAGAGGCTCATGAG 300
 Db 241 TCCACAGGTTGCGCAGGTGTTCCATCCGCGCATCCACCGCGTCCAGAGGCTCATGAG 300
 QY 301 CTGCGTGACGGTGGCGATCGCTACTTGGGCAAGGCGTGTGAAAGGCGAGTTGAAAACGTC 360
 Db 301 CTGCGTGACGGTGGCGATCGCTACTTGGGCAAGGCGTGTGAAAGGCGAGTTGAAAACGTC 360
 QY 361 AACGAAGAAATCGGCGAGAGCTCGTGGCTAGAGGCTGACGATGACGCGCTCATCGAC 420
 Db 361 AACGAAGAAATCGGCGAGAGCTCGTGGCTAGAGGCTGACGATGACGCGCTCATCGAC 420
 QY 421 GAAGCAATGATCAAGCTTGGATGGCAGCCCAACAAAGTCCCGCTGGGTGCAAGCGCAATC 480
 Db 421 GAAGCAATGATCAAGCTTGGATGGCAGCCCAACAAAGTCCCGCTGGGTGCAAGCGCAATC 480
 QY 481 CTTGGTGTTCCTATGGCTGTTGCAAGGCTGCTGCTGATTCGCGAGCGCTCCCACTGTTT 540
 Db 481 CTTGGTGTTCCTATGGCTGTTGCAAGGCTGCTGCTGATTCGCGAGCGCTCCCACTGTTT 540
 QY 541 CGCTATATCGGTGACCAACGCAACAGTCTTCCAGTTCATGATGAACATCATCAAC 600
 Db 541 CGCTATATCGGTGACCAACGCAACAGTCTTCCAGTTCATGATGAACATCATCAAC 600
 QY 601 GGTGGCGCTCAGCTGACTCGGTGTTGAGCTTCAGGAATTCATGATCGCTCCCAATCGGT 660
 Db 601 GGTGGCGCTCAGCTGACTCGGTGTTGAGCTTCAGGAATTCATGATCGCTCCCAATCGGT 660
 QY 661 GCAGAGACCTTCTCTGAGGCTTCCGCAACGGCGGAGGTTACACGCACTGAAAGTCC 720
 Db 661 GCAGAGACCTTCTCTGAGGCTTCCGCAACGGCGGAGGTTACACGCACTGAAAGTCC 720
 QY 721 GTCATCAAGAAAAGGCGCTTCCACCGGCTTGGCGATGAGGCGGCTTCGCTCCTTCC 780
 Db 721 GTCATCAAGAAAAGGCGCTTCCACCGGCTTGGCGATGAGGCGGCTTCGCTCCTTCC 780
 QY 781 GTGGCTCCACCGTGAAGGCTTGTGAGCTTATCGTTGAGGCAATCGAAGGCTGGCTTC 840
 Db 781 GTGGCTCCACCGTGAAGGCTTGTGAGCTTATCGTTGAGGCAATCGAAGGCTGGCTTC 840
 QY 841 ACCCAGGCAAGGACATCGCTTCTGCTGAGGCTTCTCTGAGTCTTCTCAAGGAC 900
 Db 841 ACCCAGGCAAGGACATCGCTTCTGCTGAGGCTTCTCTGAGTCTTCTCAAGGAC 900
 QY 901 GGCACTTACCACTTCGAAGGTGGCGACACTCCGCACTGAGATGGCAACGTTTACGCT 960
 Db 901 GGCACTTACCACTTCGAAGGTGGCGACACTCCGCACTGAGATGGCAACGTTTACGCT 960
 QY 961 GAGCTCTGTCAGGTTACCAATCGTCTCATGAGGACCCACTGCGAGGAAGATGACTGG 1020
 Db 961 GAGCTCTGTCAGGTTACCAATCGTCTCATGAGGACCCACTGCGAGGAAGATGACTGG 1020
 QY 1021 GAGGTTTACCAACCTTCAGCAACCATCGGCGAAGGTTTCAGATCGTTGGCGAGC 1080
 Db 1021 GAGGTTTACCAACCTTCAGCAACCATCGGCGAAGGTTTCAGATCGTTGGCGAGC 1080
 QY 1081 TTCTTCCTCAACCACTTCAGCGCTTGAAGGAGGCGATCGTGAAGGCTGCGCACTCC 1140
 Db 1081 TTCTTCCTCAACCACTTCAGCGCTTGAAGGAGGCGATCGTGAAGGCTGCGCACTCC 1140
 QY 1141 ATCTGTTTAAAGTTGAACAGATCGGTACCTCACCGAGACCTTCGACGCTGTCGACATG 1200
 Db 1141 ATCTGTTTAAAGTTGAACAGATCGGTACCTCACCGAGACCTTCGACGCTGTCGACATG 1200
 QY 1201 GCTCACCGCGAGGCTACACTCCATGATGCTCCACCGTTCGGTGAACCGGAGCAC 1260
 Db 1201 GCTCACCGCGAGGCTACACTCCATGATGCTCCACCGTTCGGTGAACCGGAGCAC 1260
 QY 1261 ACCATTGCTGACCTCGCAGTTGCACTCAACTGTGGCCAGATCAAGACTGTTGCTCCAGCA 1320

Db 1261 ACCATTGCTGACCTCGCAGTTGCACTCAACTGTGGCCAGATCAAGACTGTTGCTCCAGCA 1320
 QY 1321 CGTTTCGACCGTGTTCGCAAAAGTACAACAGCTTCTCCGATCGAGAGCTGCTTGGCGAC 1380
 Db 1321 CGTTTCGACCGTGTTCGCAAAAGTACAACAGCTTCTCCGATCGAGAGCTGCTTGGCGAC 1380
 QY 1381 GCCGGGCTTACGACGAGTTCGAGCGCATTCGCCAGCTTTCCAGGCTAAATAAAGCGCTT 1440
 Db 1381 GCCGGGCTTACGACGAGTTCGAGCGCATTCGCCAGCTTTCCAGGCTAAATAAAGCGCTT 1440
 QY 1441 TTGACGCCCGGTAACTCAAGGTTGCCGGGCTGCTTACTACTGTTACTGTTGTT 1500
 Db 1441 TTGACGCCCGGTAACTCAAGGTTGCCGGGCTGCTTACTACTGTTACTGTTGTT 1500
 QY 1501 GACTATGATCAGGATTTATGCGCAAGCAGAAAGAACTCATAAAGGCTTGTCTCTGCT 1560
 Db 1501 GACTATGATCAGGATTTATGCGCAAGCAGAAAGAACTCATAAAGGCTTGTCTCTGCT 1560
 QY 1561 CAAGCAGGGAACGTGCTT 1578
 Db 1561 CAAGCAGGGAACGTGCTT 1578

RESULT 3
 BD014844
 LOCUS
 DEFINITION
 Polynucleotide, DNA, isolation of polynucleotide sequence, and method of fermentative production of L-amino acid.
 ACCESSION
 BD014844
 VERSION
 JP 2001161380-A/1.
 KEYWORDS
 Corynebacterium glutamicum
 SOURCE
 Corynebacterium glutamicum
 ORGANISM
 Corynebacterium glutamicum
 Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacterineae; Corynebacteriaceae; Corynebacterium.
 REFERENCE
 1 (bases 1 to 1578)
 AUTHORS
 Meckel,V., Pfeifferle,W., Hermann,T., Puehler,A., Kalinovski,J. and Baate,B.
 TITLE
 Polynucleotide, DNA, isolation of polynucleotide sequence, and method of fermentative production of L-amino acid
 JOURNAL
 Patent: JP 2001161380-A 1 19-JUN-2001;
 DEGUSA HUELS AG
 COMMENT
 OS Corynebacterium glutamicum
 PN JP 2001161380-A/1
 PD 19-JUN-2001
 PF 04-OCT-2000 JP 2000305110
 PR 05-OCT-1999 DE 19947791.4
 PI VETTYNA MECKEL, WALTER PFEIFFERLE, THOMAS HERMANN, ALFRED PUEHLER, JOERN KALINOVSKI, BRIGITTE BAATE
 PC C12N15/09, C12N1/21, C12N9/88, C12P13/04, C12P13/08, C12N15/09, C12R1:15),
 PC C12R1:15),
 PC (C12P13/04, C12R1:15), (C12P13/08, C12R1:15), C12N15/00, (C12N15/00, PC C12R1:15)
 CC Polynucleotide, DNA, isolation of polynucleotide sequence, and method of fermentative production of L-amino acid
 CC Polynucleotide, DNA, isolation of polynucleotide sequence, and method of fermentative production of L-amino acid
 PH Key Location/Qualifiers
 FT CDS Location/Qualifiers
 FEATURES
 source
 1. 1578
 /organism="Corynebacterium glutamicum"
 /mol_type="genomic DNA"
 /db_xref="taxon:1718"
 ORIGIN
 Query Match 100.0%; Score 1578; DB 6; Length 1578;
 Best Local Similarity 100.0%; Pred. No. 6.9e-288;
 Matches 1578; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGCTGGGATATGGTAGTTTTCGCCACTAATTTCAACTGATTCCTCATCGAAACAAGA 60

```
Db 1 GGTGGGATATGGTAGTTTTCGCCACTAATTTCAACTGATTCGCTCATCGAAACAGA 60
QY 61 TTCTGCAACAAATTTGGGTAGAGCTGATTTGAAGACATTTTGAATCACTGATGAATTTCTAG 120
Db 61 TTCTGCAACAAATTTGGGTAGAGCTGATTTGAAGACATTTTGAATCACTGATGAATTTCTAG 120
QY 121 TTAGCTCCCAAGTTGGATAGGAGGCCACAGTGGCTGAATCATGACAGTATTGCTCGC 180
Db 121 TTAGCTCCCAAGTTGGATAGGAGGCCACAGTGGCTGAATCATGACAGTATTGCTCGC 180
QY 181 GAAATTCCTGACTCCCGCGGTAAACCCAAACCGTCGAGGACAGAGTTTTCCTGGATGACGGT 240
Db 181 GAAATTCCTGACTCCCGCGGTAAACCCAAACCGTCGAGGACAGAGTTTTCCTGGATGACGGT 240
QY 241 TCCACGGTTCGACAGGTTCATCCATCCGGCGCATCCACCGGCGTCCACGAGGCTCATGAG 300
Db 241 TCCACGGTTCGACAGGTTCATCCATCCGGCGCATCCACCGGCGTCCACGAGGCTCATGAG 300
QY 301 CTGCTGACGGTGGCGATCCTACCTGGGCAAGGGCGTTTGAAGGAGTTGAAACGTC 360
Db 301 CTGCTGACGGTGGCGATCCTACCTGGGCAAGGGCGTTTGAAGGAGTTGAAACGTC 360
QY 361 AACGAAGAAATCGCGACGAGCTCGCTGGCTTAGAGGCTGACGATCAGCGCCTCATCGAC 420
Db 361 AACGAAGAAATCGCGACGAGCTCGCTGGCTTAGAGGCTGACGATCAGCGCCTCATCGAC 420
QY 421 GAAGCAATGATCAAGCTTGATGGACCGCCAAACAAAGTCCCGCTGGGTGCAACGCAATC 480
Db 421 GAAGCAATGATCAAGCTTGATGGACCGCCAAACAAAGTCCCGCTGGGTGCAACGCAATC 480
QY 481 CTTGGTGTTCCTATGGCTGTTCGAAAGGCTGCTGATTCGCGAGGCTCCGACTGTTTC 540
Db 481 CTTGGTGTTCCTATGGCTGTTCGAAAGGCTGCTGATTCGCGAGGCTCCGACTGTTTC 540
QY 541 CGCTACATCGGTGGACAAACGACACAGTTCCTCCAGTTCCTCAATGATGAACATCATCAAC 600
Db 541 CGCTACATCGGTGGACAAACGACACAGTTCCTCCAGTTCCTCAATGATGAACATCATCAAC 600
QY 601 GGTGGCGCTCAGCTGATCGGTGTTGAGCTTCAGGAATTCATGATCGTCCAAATCGGT 660
Db 601 GGTGGCGCTCAGCTGATCGGTGTTGAGCTTCAGGAATTCATGATCGTCCAAATCGGT 660
QY 661 GCAGAGACCTTCTCTGAGGCTCTCCGCAACGGCGGAGTCTACCAACGACCTGAAAGTCC 720
Db 661 GCAGAGACCTTCTCTGAGGCTCTCCGCAACGGCGGAGTCTACCAACGACCTGAAAGTCC 720
QY 721 GTCATCAAGGAAAGGGCCTGTCCACCGGACTTGGCGATGAGGGCGGCTTCGCTCCCTTC 780
Db 721 GTCATCAAGGAAAGGGCCTGTCCACCGGACTTGGCGATGAGGGCGGCTTCGCTCCCTTC 780
QY 781 GTGGCTCCACCGTGAAGGCTCTTGACCTTATCGTTGAGGCAATCGAGAAAGGCTGGCTTC 840
Db 781 GTGGCTCCACCGTGAAGGCTCTTGACCTTATCGTTGAGGCAATCGAGAAAGGCTGGCTTC 840
QY 841 ACCCGAGGCAAGACATCGCTCTTGCTCTGAGGCTTCTCTGAGTTCCTCAAGGAC 900
Db 841 ACCCGAGGCAAGACATCGCTCTTGCTCTGAGGCTTCTCTGAGTTCCTCAAGGAC 900
QY 901 GGCACCTTACCCTTCGAAGGTGGCGACGACTCCGCGAGCTGAGATGGCAACGTTTACGCT 960
Db 901 GGCACCTTACCCTTCGAAGGTGGCGACGACTCCGCGAGCTGAGATGGCAACGTTTACGCT 960
QY 961 GAGCTCGTTGACGGTACCAATCGTCTCATCGAGACCCACTGACGGAAGATGACTGG 1020
Db 961 GAGCTCGTTGACGGTACCAATCGTCTCATCGAGACCCACTGACGGAAGATGACTGG 1020
QY 1021 GAGGTTTACACCAACCTCACCGCAACCATCGGCGACAAGGTTTCAGATCGTTGGCGAC 1080
Db 1021 GAGGTTTACACCAACCTCACCGCAACCATCGGCGACAAGGTTTCAGATCGTTGGCGAC 1080
QY 1081 TTCTTGTCTACCAACCTGAGCGCTGGAAGGAGGCAATCGCTAAGAGGCTGCCAACTCC 1140
Db 1081 TTCTTGTCTACCAACCTGAGCGCTGGAAGGAGGCAATCGCTAAGAGGCTGCCAACTCC 1140
```

```
Db 1081 TTCTTGTCTACCAACCTGAGCGCTGGAAGGAGGCAATCGCTAAGAGGCTGCCAACTCC 1140
QY 1141 ATCTCGTTAAGGTGAACAGATCGGTACCTTCCACGAGACCTTCGACGCTGTCGACATG 1200
Db 1141 ATCTCGTTAAGGTGAACAGATCGGTACCTTCCACGAGACCTTCGACGCTGTCGACATG 1200
QY 1201 GCTCACCGCGAGGCTACACCTCCATGATGTCCACCGTTCGGTTCGAGACCGGAGCAC 1260
Db 1201 GCTCACCGCGAGGCTACACCTCCATGATGTCCACCGTTCGGTTCGAGACCGGAGCAC 1260
QY 1261 ACCATTGCTGACCTCCGAGTTGACATCAACTGTGGCGAGATCAAGACTGTGCTCCAGCA 1320
Db 1261 ACCATTGCTGACCTCCGAGTTGACATCAACTGTGGCGAGATCAAGACTGTGCTCCAGCA 1320
QY 1321 CGTTCCGACCGGTTCGCAAAAGTAAACACAGCTTCTCCGATCGAGAGCTGCTTGGCGAC 1380
Db 1321 CGTTCCGACCGGTTCGCAAAAGTAAACACAGCTTCTCCGATCGAGAGCTGCTTGGCGAC 1380
QY 1381 GCGGCGCTTACGACAGGTGCGACGCTTCCACGCTTTCAGGGCTAAATATAAGGCGCTT 1440
Db 1381 GCGGCGCTTACGACAGGTGCGACGCTTCCACGCTTTCAGGGCTAAATATAAGGCGCTT 1440
QY 1441 TTCCGACCGCGGTAACTCAAGTTGCGCGGCGTCTGCTTACTTACTGTTACTGTTGT 1500
Db 1441 TTCCGACCGCGGTAACTCAAGTTGCGCGGCGTCTGCTTACTTACTGTTACTGTTGT 1500
QY 1501 GACTATGATCGAGGATTTATGGCAAGAGAGAGAGAACTATATAAGGCTTGTTCCTGTCT 1560
Db 1501 GACTATGATCGAGGATTTATGGCAAGAGAGAGAGAACTATATAAGGCTTGTTCCTGTCT 1560
QY 1561 CAAGCAGGGAACGTGCTT 1578
Db 1561 CAAGCAGGGAACGTGCTT 1578

RESULT 4
AP005277
LOCUS
DEFINITION
AP005277 Corynebacterium glutamicum ATCC 13032 DNA, linear BCT 08-AUG-2002
4/10.
ACCESSION
VERSION AP005277 BA000036
KEYWORDS AP005277.1 GI:21323710
SOURCE
ORGANISM Corynebacterium glutamicum ATCC 13032
Corynebacterium glutamicum ATCC 13032
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Corynebacteriaceae; Corynebacterium.
REFERENCE
1 Nakagawa, S.
Complete genomic sequence of Corynebacterium glutamicum ATCC 13032
Unpublished
2 (bases 1 to 333150)
Direct Submission
Submitted (24-MAY-2002) Satoshi Nakagawa, Kyowa Hakko Kogyo Co.
Ltd., Tokyo Research Laboratories; 3-6-6, Asahi-machi, Machida,
Tokyo 194-8533, Japan (E-mail:snakagawa@kxanagen.com,
Tel:81-44-829-3031, Fax:81-44-813-1651)
This sequence is conducted by collaboration of Kyowa Hakko Kogyo
Co. Ltd. And Kitasato University.
FEATURES
Location/Qualifiers
1..333150
/organism="Corynebacterium glutamicum ATCC 13032"
/mol_type="genomic DNA"
/strain="ATCC 13032"
/db_xref="taxon:196627"
/notes="ATCC 13032"
/complement(63..1292)
/gene="Cgl0944"
/complement(63..1292)
/gene="Cgl0944"
/codon_start=1
/transl_table=11
gene
CDS
```

```

/product="Hypothetical membrane protein"
/protein_id="BAB98337.1"
/db_xref="GI:21323711"
/translation="MTNPTEERNARRLIWANGLQNIQDOIVAAKTVLPWLLQAAAGP
FLALLVPIRAGSMLOAAITGWVLQTSRKVVGISGQFVSALGIGVAAFLRAG
WALGITVILVLAALSLPRSCGSIASKOVQKVISKGRGLVTGRATVIGVMGLVAGL
AIATFLGSHSTRVLAUVIASSFSWLPASIFARIEYAKPATPKNAPSANPVRRCI
AALKDKAFRRFVLVRSMVLTAFTAPIVALAAESGNSIDSLGFFLIASGLASMVG
RISGIDSHSKNMGAGLFGSIVLILVLSAPAQNTLVPLPSFFLITLAHTA
IRVARKYVMDAEGDQRTYRVADANTLMGVLLIVGALSQFIAPGNEALLFLAAI
GLICTISARGLKEVSAG"
1382..2917
/gene="Cg10945"
1382..2917
/gene="Cg10945"
/codon_start=1
/transl_table=11
/product="Putative multicopper oxidases"
/protein_id="BAB98338.1"
/db_xref="GI:21323712"
/translation="MPELNRRTFFKGAQVLAATVVGQVLVACSSDDVRVGYGEPRTL
PIPADLGTREGSVHFALAEQTCESQILPDVTTKWTGNGTHLGPVLVVKGGDDVHV
DVINLDEMVTVMHGMKLPALADGPHSPIGPQWSPPTWTVANDAATLWYRPHTHG
LTGCHAYRGLAGLIIIDEADTKDLPRFYGVDIPLVMDHREFDGSDEEDLPDL
GLIGDTTANGITNAHFDAITRRVRFRVNGSNMRFYNLAFSDTRTFQVIASSGLLD
EPQDRTTLAIGGERWEIVVEFEGEDVTLESVGEDNYGVPDDFVDFCMDSFOL
LTTITGSDDAAPALPGVLVKSTEPDVIDATERTFTMNTFSINDLQMDQRVDDVID
HDQEVMIVTNDSMDPHNFVHDAREFKVLKFECDVELEFDGKQVTVGLPGCATL
AVEGCHYDPQWPYMYCHMLYHEDQGMQGFVIVEPGEPAAVLGSCTGSSIDSAGG
HAH"
3063..3983
/gene="Cg10946"
3063..3983
/gene="Cg10946"
/notes="PF00005:ABC transporter
TIGR00960:3a0501802: Type II (General) Secretory Pathway
(IIS) Family protein
TIGR00968:3a0106801: sulfate transport system permease
protein
TIGR01166:cbio: cobalt transport protein ATP-binding
subunit
TIGR01184:ntrCD: nitrate transport ATP-binding subunits C
and D
TIGR01186:prov: glycine betaine/L-proline transport ATP
binding subunit
TIGR01187:potA: spermidine/putrescine ABC transporter
ATP-binding subunit
TIGR01188:dra: daunorubicin resistance ABC transporter
ATP-binding subunit
TIGR01189:ccmA: heme exporter protein Ccma"
/codon_start=1
/transl_table=11
/product="ABC-type transporter, ATPase component"
/protein_id="BAB98339.1"
/db_xref="GI:21323713"
/translation="MTTPAIDVTLVRYTGDYTAVKGLNFHVORGFVGLIGTNGAG
KTSLEVEGLSAPSSCTVRIISGLDPVADRAILPELIMLQSGLPSPQTVTAETMDM
WHTCTYPRAIKQVLDVLLHRENVKVGALSGBEQRRLACALIGDPSILFDEPT
LGDPERRHTWQLLLKQGVMTMLTTHYLEAEFLCDRIAIMNAGEIAVGTGLDE
LVAREKSIISFLRGQVPELPSLGAIEIIRDNNHVRTATTLOOHTLEILTAAETGI
ALGFAAKPATLESVFMDIASLENTSLQTA"
4003..4836
/gene="Cg10947"
4003..4836
/gene="Cg10947"
/codon_start=1
/transl_table=11
/product="ABC-type transporter, permease components"
/protein_id="BAB98340.1"
/db_xref="GI:21323714"
/translation="MTTSHTRAGLQHAHPERRKTSFFKTSFLKAEMLQFRNKTLFLFM
ATVPVGIPLLLFLINGGAESANSFDFYMTLLFVQFYTVLSMATTRDRDLVKR
LRTGEARDIDIIGAICPGALLTLFTWIIPLLMVLGAPAPINLPIVFAVLIGLL

```

```

CSALALMTSGFTRNAEAAQMTSMPVFVFLMAMGGLGSIIRFVFGDSIVADILAYTPFAAIS
DLVQIWAGATFADSVGVEAANFAGIFQDMLIPLGLIAATAAANAANRMRWDSY
R"
4879..6138
/gene="Cg10948"
4879..6138
/gene="Cg10948"
/EC number="2.7.3.-"
/codon_start=1
/transl_table=11
/product="Two-component system, sensory transduction
histidine kinases"
/protein_id="BAB98341.1"
/db_xref="GI:21323715"
/translation="MNKDFWTAGTWARTARWFRSGVSLASPVTPAPLNSWRRLLPNLAKYTL
YTRVSLQAPVLLSAYFLGIVANAGTLNPSFVWLGLFSVILLIIVTVLVVEYQPSLNS
YPRSVQPPFFTGVLVNLGVVSVVLIQIFGLNMSDNTRATALLFTLTCVLLSIAYI
PMWYRWLIAMSVLAWMTSTTIDYLSALWVVIPLPMAGTVRLSVTVDMKVEVRS
RELEASLRVTEERLRFQAEHLDTLGHQAAMSVKSELAALAKRGDDRLNELELQK
LRTSMEKRDVSVGYRTVNLAITEIEGAKSILLADAHILHSVIGTTSQVSPAHRELCAW
LVREATTNLRHSDATDATLTLTSSTEVMDNNGVNDIGRLSGLSALRSRAESAGMTL
IVSREDDQFSVRMLINAPANTPAEKA"
6139..6750
/gene="Cg10949"
6139..6750
/gene="Cg10949"
/notes="PF00072:Response regulator receiver domain
PF00196:Bacterial regulatory proteins, luxR family"
/codon_start=1
/transl_table=11
/product="Two-component system, response regulators
consisting of a CheY-like receiver domain and a HTH
DNA-binding domain"
/protein_id="BAB98342.1"
/db_xref="GI:21323716"
/translation="MISISIADEALIASLATLLSLEPDLVDPRTAGSGEELIETWA
DPNRTDVCVDLQIGDIGIDITRLMETTPDLAVITVSHARPQIKRALAAGVLG
FLKTSADAEFATAIRTVHAGRVYIDPELAAMTISAGESPLTNEEEVLELAGQGLSA
EETAAHAPGTRNYLSOAMTKVGAQNRFEATFRARELGWL"
complement(6781..6987)
/gene="Cg10950"
complement(6781..6987)
/gene="Cg10950"
/codon_start=1
/transl_table=11
/product="Hypothetical protein"
/protein_id="BAB98343.1"
/db_xref="GI:21323717"
/translation="MKHYWSMYFGAGPLNTPHTKEVMFMDIVSIVNHFWEFWSSTNNWVN
MTGLPRETFVPAFVGSVTLVLSF"
7048..8205
/gene="Cg10951"
7048..8205
/gene="Cg10951"
/codon_start=1
/transl_table=11
/product="tRNA-processing ribonuclease BN"
/protein_id="BAB98344.1"
/db_xref="GI:21323718"
/translation="MTAWLCSFPSPQVDEYDMNGVVQPOEHLDATLIAADFHGPNP
SGRKLRLNPOGKVALNTRVDVDPDGLDLAALLTFFSILSIAPAVLLGYSVITIF
LASDTEILNLVDEVNVQVYVPEQSHVNVGVIDSIAGSAAAGOVGAVGVITALTWS
AYVAFSRCANAVYGRSEKTLTKRWMLLFLNALLLLGIIILVSVWLNELVMGILF
APIAEPLHLTNVLSFLITDRPMPIWIWRFPVIVGVLMFVATLYWAPNARPMKFRWL
SLGSFLAIVGILLAGVNFYFTLFAAFSSYGAAGVLSLAVFIALWVFNICLIITGLKID
VEISRAKQLQAGPAEDYSLVPPRSIEKVAKMQRQORLMDQAAARESN"
8672..10114
/gene="Cg10952"
8672..10114
/gene="Cg10952"
/notes="PF00005:ABC transporter
PF00664:ABC transporter transmembrane region."
/codon_start=1

```

```
/transl_table=11
/product=ABC-type multidrug/protein/lipid transport
/system, AtPase component"
/protein_id="BAB98345.1"
/db_xref="GI:21323719"
/translations="MRTTHAGSILRRMIRROKGVAFGLGMWQLSEALVPIAI
GLIVDHAVLTKDLRLVGLVAVPVLFVLSFSYRFGSRALNRVNPESHAREVAD
HAUKNDLPRNLVEGVMSRSTADSDTSIFGQIGTGVSAAATFLGAATYLLISDWLV
GLIVLVLPIIISGVVLAASKGISKRSVTQOEKLAESGAQASDIMGLRVITKATGGBRW
AVTFEKVASQASARAADTAVASGVAGIGELSIYVNLAAVLLAGWRVTTGELGPRQ
LIAIVGVAVLSPRIILSNSINASAIHAAGAEVANFLNLDSEQAQYSESSEINDGE
FLIVVPASTLPKGDNLATPHADLIFEGTLRSNI SMNHEDNVPIDPOVIRASGLTDI
IEVDGLDAPRDTGSNLSGQORVALARALHADAEVLVLMDFTSADVSTVEISIAQG
IKQURAGKTLIVSSSPAFYNLADRIVISHV"
/gene 10107..11843
CDS /gene="Cg10953"
10107..11843
/gene="Cg10953"
```

```
Query Match 100.0%; Score 1578; DB 1; Length 333150;
Best Local Similarity 100.0%; Pred. No. 4.9e-288;
Matches 1578; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGCTGGGGATATGGTAGTTTTCGCCCACTAATTTCAACTGATTGCCCTCATCGAAACAAGA 60
Db 34849 GGCTGGGGATATGGTAGTTTTCGCCCACTAATTTCAACTGATTGCCCTCATCGAAACAAGA 34908

Qy 61 TTCGTGCAACAATTTGGGTAGACGTGATTGAAGACATTTTGAACGTCGAATTAATCTAG 120
Db 34909 TTCGTGCAACAATTTGGGTAGACGTGATTGAAGACATTTGATCACTGGAATTAATCTAG 34968

Qy 121 TTAGCTCCCAAGTTGGCATAGAGGCCACAGTGGCTGAAATCATGCAAGTATTGCTCGC 180
Db 34969 TTAGCTCCCAAGTTGGCATAGAGGCCACAGTGGCTGAAATCATGCAAGTATTGCTCGC 35028

Qy 181 GAAATTTCTGACTCCCGCGGTAAACCAACCGTCGAGGACAGGTTTTCCTGGATGACGGT 240
Db 35029 GAAATTTCTGACTCCCGCGGTAAACCAACCGTCGAGGACAGGTTTTCCTGGATGACGGT 35088

Qy 241 TCCACAGGTGTCGAGGTGTTTCCATCCGGGCGATCCACCGCGCTCCACGAGGCTCATGAG 300
Db 35089 TCCACAGGTGTCGAGGTGTTTCCATCCGGGCGATCCACCGCGCTCCACGAGGCTCATGAG 35148

Qy 301 CTGGGTGACGGTGCGGATCGTACTCTGGGCAAGGGCGTTTGAAGGACGATTGAAAACGTC 360
Db 35149 CTGGGTGACGGTGCGGATCGTACTCTGGGCAAGGGCGTTTGAAGGACGATTGAAAACGTC 35208

Qy 361 AACGAAGAAATCGGCGACGAGCTCGCTGGCTAGAGCTGACGATCAGCGCTCATCGAC 420
Db 35209 AACGAAGAAATCGGCGACGAGCTCGCTGGCTAGAGCTGACGATCAGCGCTCATCGAC 35268

Qy 421 GAAGCAATGATCAAGCTTGATGGCACCGCCCAACAAAGTCCCGCTCGGTGCAACGCAATC 480
Db 35269 GAAGCAATGATCAAGCTTGATGGCACCGCCCAACAAAGTCCCGCTCGGTGCAACGCAATC 35328

Qy 481 CTTGGGTGTTCCATGGCTGTTGCAAGGCTGCTGATTCGGCAGGCGCTCCCACTGTT 540
Db 35329 CTTGGGTGTTCCATGGCTGTTGCAAGGCTGCTGATTCGGCAGGCGCTCCCACTGTT 35388

Qy 541 CGCTACATCGGTGACCAACGACACGTTCTTCCAGTTTCCAATGATGAACATCATCAAC 600
Db 35389 CGCTACATCGGTGACCAACGACACGTTCTTCCAGTTTCCAATGATGAACATCATCAAC 35448

Qy 601 GGTGGCGCTCACGCTGACTCGGTTGTTGACGTTTCAGGAATTCATGATCGCTCCCAATCGGT 660
Db 35449 GGTGGCGCTCACGCTGACTCGGTTGTTGACGTTTCAGGAATTCATGATCGCTCCCAATCGGT 35508

Qy 661 GCAGAGACCTTCTCTGAGGCTCTCGCAACGGCGGAGGCTTACCACGACTGAAATCC 720
Db 35509 GCAGAGACCTTCTCTGAGGCTCTCGCAACGGCGGAGGCTTACCACGACTGAAATCC 35568

Qy 721 GTCATCAAGGAAAAGGCGCTGTCACCGGACTTGGCGATGAGGGCGGCTTCGCTCCTTCC 780
```

```
Db 35569 GTCATCAAGGAAAAGGCGCTGTCACCGGACTTTGGCGATGAGGGCGGCTTCGCTCCTTCC 35628
Qy 781 GTCGGGTCCACCGGTGAGGCTCTTGACCTTATCGTTGAGGCAATCGAGAAAGGCTTGGCTTC 840
Db 35629 GTCGGGTCCACCGGTGAGGCTCTTGACCTTATCGTTGAGGCAATCGAGAAAGGCTTGGCTTC 35688
Qy 841 ACCCAGGCAAGGACATCGCTCTTGCTCTGGAGCTTTCCTCTGAGTTCTTTCAGGAC 900
Db 35689 ACCCAGGCAAGGACATCGCTCTTGCTCTGGAGCTTTCCTCTGAGTTCTTTCAGGAC 35748
Qy 901 GGCACCTTACCACATTCGAAGGTGCCAGCACTCCGCACTGAGATGGCAACGTTTACGCT 960
Db 35749 GGCACCTTACCACATTCGAAGGTGCCAGCACTCCGCACTGAGATGGCAACGTTTACGCT 35808
Qy 961 GAGCTCGTTGACCGGTACCCAAATCGTCTCATCGAGACCCACTGTCAGGAAGATGACTGG 1020
Db 35809 GAGCTCGTTGACCGGTACCCAAATCGTCTCATCGAGACCCACTGTCAGGAAGATGACTGG 35868
Qy 1021 GAGGTTTACACCAACCTCACCGCAACCTCGGCAAGGTTTCAGATCGTTGGCGAC 1080
Db 35869 GAGGTTTACACCAACCTCACCGCAACCTCGGCAAGGTTTCAGATCGTTGGCGAC 35928
Qy 1081 TTCTTCTGTCACCAACCTCGAGCGCTGAAGGAGGCGATCGCTAAGAGGCTGCAACTCC 1140
Db 35929 TTCTTCTGTCACCAACCTCGAGCGCTGAAGGAGGCGATCGCTAAGAGGCTGCAACTCC 35988
Qy 1141 ATCTCTGTTAAGGTGAACAGATCGGTACCTCACCGAGACCTTTCGACGCTGTCGACATG 1200
Db 35989 ATCTCTGTTAAGGTGAACAGATCGGTACCTCACCGAGACCTTTCGACGCTGTCGACATG 36048
Qy 1201 GCTCACGGCGAGGCTACACCTCCATGATGTCCACCGTTCCGGTGAGACCGGAGCAC 1260
Db 36049 GCTCACGGCGAGGCTACACCTCCATGATGTCCACCGTTCCGGTGAGACCGGAGCAC 36108
Qy 1261 ACCATTGCTGACCTCGCAGTTGCACTCAACTGTGGGCAAGATCAAGACTGTGTGCTCCAGCA 1320
Db 36109 ACCATTGCTGACCTCGCAGTTGCACTCAACTGTGGGCAAGATCAAGACTGTGTGCTCCAGCA 36168
Qy 1321 CGTTCCGACCGGTGCGAAAGTACACAGCTTCTCCGATCGAGCAGCTGCTTGGCGAC 1380
Db 36169 CGTTCCGACCGGTGCGAAAGTACACAGCTTCTCCGATCGAGCAGCTGCTTGGCGAC 36228
Qy 1381 GCCGGGCTCTACCGAGGTGCGAGGATTCACCGCTTTCAGGCTTAAATAAAGCGCTT 1440
Db 36229 GCCGGGCTCTACCGAGGTGCGAGGATTCACCGCTTTCAGGCTTAAATAAAGCGCTT 36288
Qy 1441 TTCGACGCGCGGTAACTCAAGGTTGCGGGCGCTCGTTGCTTACTCTGTTACTGTGT 1500
Db 36289 TTCGACGCGCGGTAACTCAAGGTTGCGGGCGCTCGTTGCTTACTCTGTTACTGTGT 36348
Qy 1501 GACTATGATCGAGGATTTGCGAAAGAGAGAAACTCATAAAGGCTTGTTCCTGTCT 1560
Db 36349 GACTATGATCGAGGATTTGCGAAAGAGAGAAACTCATAAAGGCTTGTTCCTGTCT 36408
Qy 1561 CAAGCAGGGAACCTGCTT 1578
Db 36409 CAAGCAGGGAACCTGCTT 36426
```

```
RESULT 5
BX927150
LOCUS BX927150 348475 bp DNA linear BCT 10-JUN-2004
DEFINITION Corynebacterium glutamicum ATCC 13032, IS fingerprint type 4-5,
complete genome; segment 3/10.
ACCESSION BX927150 BX927147
VERSION BX927150.1 GI:41324904
KEYWORDS complete genome.
SOURCE Corynebacterium glutamicum ATCC 13032
ORGANISM Corynebacterium glutamicum ATCC 13032
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.
REFERENCE 1 (bases 1 to 348475)
AUTHORS Kalinowski,J., Bathe,B., Bartels,D., Biscoff,N., Bott,M.,
```


Burkovski, A., Dusch, N., Eggeling, L., Eikmanns, B. J., Gaigalat, L., Goesmann, A., Hartmann, M., Huthmacher, K., Kramer, R., Linke, B., McHardy, A. C., Meyer, F., Mockel, B., Pfeifferle, W., Puhler, A., Rev, D. A., Ruckert, C., Rupp, O., Sahm, H., Wendisch, V. F., Wiegand, I. and Tauch, A.

The complete *Corynebacterium glutamicum* ATCC 13032 genome sequence and its impact on the production of L-aspartate-derived amino acids and vitamins

J. Biotechnol. 104 (1-3), 5-25 (2003)

22830012

12948626

2 (bases 1 to 348475)

Kalinowski, J.

Direct Submission

Submitted (21-JAN-2004) Joern Kalinowski, Institut fuer Genomforschung, Universitaet Bielefeld; Universitaetstrasse 25, 33615 Bielefeld, Germany

E-mail: Joern.Kalinowski@cbitec.uni-bielefeld.de

This sequence was accomplished by collaboration between Degussa AG and Bielefeld University.

Join(BX927148.1:1..348071, BX927149.1:1:51..349887, BX927150.1:1:51..348475, BX927152.1:1:51..349799, BX927153.1:1:51..349584, BX927154.1:1:51..349575, BX927155.1:1:51..349136, BX927156.1:1:51..349115, BX927157.1:1:51..140057).

Location/Qualifiers

1..348475

/organism="Corynebacterium glutamicum ATCC 13032"

/mol_type="Genomic DNA"

/strain="DSM 20300 = ATCC 13032"

/db_xref="taxon:196627"

/note="IS fingerprint type: 4-5"

complement(381..1610)

/locus_tag="cg0783"

complement(381..1610)

/locus_tag="cg0783"

/locus_start=1

/codon_start=1

/transl_table=11

/product="conserved hypothetical protein"

/protein_id="CAF19386.1"

/db_xref="GI:41324905"

/translation="MAEKGVVVKLSLSEGFVTLWAPNREGNQAFGLGANDDLY VFNSELLVFESDDKDLTSHPEWKKFSGDAARVVPDANAELDIIAGALLADRP SYANVTLRIFFAMRSLGHVTAASPVTVFSSHSVLSVNDRGSEHYSGNGLGWTS VGRAVATNMGVVADLDEALTVEKNSLFTEDATRIQEAQAQVQAAAEAEQIAKEE AEKVDYDNSPWAAGIDPIKVIDGRTIYTLRYLGGQVFLKFGKEIFTSQKSL LRWLVDHDDHLARASTSDMLGINAGETELLVHSNVSFNGLVKDIINTSDVADTS QMARAYELMADTADWADDSDMSNYFLANPRMDIYISYMLGSDTSGYVPHAFNDHS ESWELEEMLIKRFKSF"

complement(1677..2537)

/locus_tag="cg0784"

complement(1677..2537)

/locus_tag="cg0784"

/locus_start=1

/codon_start=1

/transl_table=11

/product="PUTATIVE CELL WALL-ASSOCIATED HYDROLASE"

/protein_id="CAF19387.1"

/db_xref="GI:41324906"

/translation="MIGLISALSRLTHEPAQLPOLAIPQAPDISAAILGRELSANP TTLISLFDATSOHLHIVTSLIOAAPLIDAAHDIAATQAHLNQAAGLVMRFSWNP AESFAARTELSLPLGLMIAEASORLEMSRQVPIVEKLVOIQDLPTEPTAEVETA PVTMAQAQPSNDNEVGKQVAQAKEALGTPYLMGTSSTGFCDSGLTQAWRAGVEI PRIADQAVGRQVYBELQEGDLLINDGHVAMTAGGQIIIEAGDPVOLNPIRTSNIGM AFHGYRFRPT"

complement(2534..2797)

/locus_tag="cg0785"

complement(2534..2797)

/locus_tag="cg0785"

/locus_start=1

/transl_table=11

/product="hypothetical protein predicted by Glimmer/Critica"

/protein_id="CAF19388.1"

/db_xref="GI:41324907"

/translation="MRINLPHAKELAHLCILPTPAVPALPTDSCAOFIDHQAISASI ATVALNTLSHTAENLGNALTCGLAEIETDQLAHALERLT"

2896..3531

/gene="upp"

/locus_tag="cg0786"

2896..3531

/gene="upp"

/locus_tag="cg0786"

/EC_number="2.4.2.9"

/codon_start=1

/transl_table=11

/product="PUTATIVE URACIL PHOSPHORIBOSYLTRANSFERASE"

/protein_id="CAF19389.1"

/db_xref="GI:41324908"

/translation="MDITIVNPLVASRLTLRDRSDNAAFRAAANDLGAMLIYEAS RDLVEHPDFTKTPVMAEGTRLKQPPILVPIIRAGLGMIDPALSMDAQQVGFGLGAR DEETHFPVLEALPQDLSNQPVFLVDPMLATGSLHAIIRLLADRATDITATCMVS AQPQVDALAESGLPVLVATIIDPLGDENAYIVPGLDAGDRLYGPRNIDL"

3625..3999

/locus_tag="cg0787"

3625..3999

/locus_tag="cg0787"

/codon_start=1

/transl_table=11

/product="transcriptional regulator"

/protein_id="CAF19390.1"

/db_xref="GI:41324909"

/translation="MSEKRWSSYGHAFASRLKRLTURGFSQEBELADLSGVSRTNIS YERNENKGNAPDQSLNRYRLAQALDVPPIALMPAGSVPAKICVDETAIDVRWPS EKDPILFDADLRLTRNREQPN"

complement(3996..5588)

/gene="pmmB"

/locus_tag="cg0788"

complement(3996..5588)

/gene="pmmB"

/locus_tag="cg0788"

/EC_number="5.4.2.2"

/EC_number="5.4.2.8"

/codon_start=1

/transl_table=11

/product="PHOSPHOGLUCOMUTASE/PHOSPHOMANNOMUTASE"

/protein_id="CAF19391.1"

/db_xref="GI:41324910"

/translation="MDESRLSFGTAGRAPVGPABHOMNVQVTRTAGVASWLAER AALNPHLVPEDETIGIRALYFDQGLRVVVDARYGSHTRFAATTAETVAGAFEV TLLPTSPPTLIPLVKNHGLDAGVQITASHGAADNGYKVLNSGRQLYSELEPELE AHINAVEDPTIRVPRVTRPTADQLRRYVDEMVSILVTPQADLLRVNSRGNLRVVYTA LHGVGGRMANAFQFAGFPHTHGVKAQYQDPPTPTVAFNPPEPSAIELLLERAKEK NADILFALDPDADRCVAGIRTADGHRMLSGDEVTLLATRLVPEYSGEGPRPVATT VVSSQLLIIAEDKGDYSETLIGFKNLSRAADGLDGLPLAFAYEEAVGTCTPVPDVPD XDGISTALFNASWAAELQAQSLQKLNLYRRYGYFASQIATVTSSTPRELVDHWI AHPQQLIGVSVTPHILPEKQIALHGVGHVHTRAIATRVSGTGEAKRLYLEVQQAAS HDEAAQLHQLEDEVQSWLSKL"

5693..6880

/gene="amiA"

/locus_tag="cg0789"

5693..6880

/gene="amiA"

/locus_tag="cg0789"

/codon_start=1

/transl_table=11

/product="PUTATIVE N-ACYL-L-AMINO ACID AMIDOHYDROLASE"

/protein_id="CAF19392.1"

/db_xref="GI:41324911"

/translation="MMEIGVQVASMWRHDEVIKWRRLHSHPELSHMEYRTTEYLA SVLKDHGMREHLPGTGLMVDIGEGDSRLAFRADIDALPILSTGTEFSSTATGVAAH ACCHDYHTVIALALACALNTIELPIGIRVIFQPAEEVMTGATDVAHGLDGVDAIY AIHVEPKLVKRVGVRAGAITASDVIRKVGEGHSARPHUSADVVYALSKLVVDL PGLLSRVRDPTGTIYVFGTINAGYAPNAIPDSGIVSGTLRTADITWDRMRLISEL VEQLAPTGVTHIELIYVPGVPPVINDDDATALLASARAMDMDTDSVQPOSSGGEDFS WYLEHVPGSMARLGCWPGHGPKQDLHQSLVDVBERAIGVGVRLFGSLVQVYSSRSEAF LNS"

gene	7020. .8429	Db	338771	AACGAGAAATCGGCGACGAGCTCGCTGGCTAGAGCTGACGATCAGCGCCTCATCGAC	338830
CDS	/gene="lpda"	Qy	421	GAAGCAATGATCAAGCTTGAATGGCAACGCAACAAAGTCCCGCTGGTGAAGCGCAATC	480
	7020. .8429	Db	338831	GAAGCAATGATCAAGCTTGAATGGCAACGCAACAAAGTCCCGCTGGTGAAGCGCAATC	338890
	/gene="lpda"	Qy	481	CTTGGTGTTCATGCTGTGTTGCAAAAGGCTGTGCTGATTCCCGCAGGCTCCCACTGTC	540
	/locus_tag="cg0790"	Db	338891	CTTGGTGTTCATGCTGTGTTGCAAAAGGCTGTGCTGATTCCCGCAGGCTCCCACTGTC	338950
	/EC_number="1.8.1.4"	Qy	541	CGCTACATCGGTGGACCAACGACACGCTTCTCCAGTTCCTCAATGATGAACATCATCAAC	600
	/codon_start=1	Db	338951	CGCTACATCGGTGGACCAACGACACGCTTCTCCAGTTCCTCAATGATGAACATCATCAAC	339010
	/transl_table=11	Qy	601	GGTGGCGCTCACGCTGACTCCGCTGTGAGCTTCAGGAATTCATGATCGCTCCCAATCGGT	660
	/product="DIHYDROLIPOAMIDE DEHYDROGENASE"	Db	339011	GGTGGCGCTCACGCTGACTCCGCTGTGAGCTTCAGGAATTCATGATCGCTCCCAATCGGT	339070
	/protein_id="CAF19393.1"	Qy	661	GCAGAGACCTTCTCTGAGGCTCTCCGCAACGCGCGGAGGTCTACACGCACTGAAGTCC	720
	/db_xref="GI:41324912"	Db	339071	GCAGAGACCTTCTCTGAGGCTCTCCGCAACGCGCGGAGGTCTACACGCACTGAAGTCC	339130
	/translations="MAKRIVIIGGSPAGYEAALAGAKYGAETVVIDVGVGSAVTMD	Qy	721	GTCTCAAGGAAAGGCGCTTGTCCACCGGACTTTGGCGATGAGGCGCGCTTCGCTCTCTCC	780
	CVPKSGFIAGTKTDLRADDMLNRLGKHALEIDALNIRVKDLAKAQSEDLGOL	Db	339131	GTCTCAAGGAAAGGCGCTTGTCCACCGGACTTTGGCGATGAGGCGCGCTTCGCTCTCTCC	339190
	QRSVDMINGVGFEDYNTKQTHYIKVTHSDSESTVECDLVLVATGATPRLLKAE	Qy	781	GTCCGCTCCACCGCTGAGGCTCTTGACCTTATCGTTGAGCAATCAGAAAGGCTGGCTTC	840
	PDGERLLTRQVTDIELPHTLLVVGSGTVGAFSAFAGELGVKVTVAASRDRLPHD	Db	339191	GTCCGCTCCACCGCTGAGGCTCTTGACCTTATCGTTGAGCAATCAGAAAGGCTGGCTTC	339250
	DADAADVLETVAERGVSLEKHARVESVTRTDDGGVCRTADREIYGSHALMTGSI	Qy	841	ACCCAGGCAAGGACATCGCTTGTCTGAGGCTTGTCTCTCTGAGTTCCTGAGTTCCTCAAG	900
	PNTADLGNIGVLAAPSIIKVDVRSRTNIPGYAAGDCTDLFFPLASVAAAMQGRATM	Db	339251	ACCCAGGCAAGGACATCGCTTGTCTGAGGCTTGTCTCTCTGAGTTCCTGAGTTCCTCAAG	339310
	YHALGVSVPRLUKTVATVTRPEIAAAGITHAQVDSGVSARVTVLPLATNPRAM	Qy	901	GGCACTTACACCTTCGAAGGTGCCAGCACTCCGCACTGAGATGGAAGCGTTTACGCT	960
	YSLRHGFVKLFCRNSGLLIIGVVVAPTASELILPIAVAVTNRLTVADLADTFAVYPS	Db	339311	GGCACTTACACCTTCGAAGGTGCCAGCACTCCGCACTGAGATGGAAGCGTTTACGCT	339370
	LSGSITFAARQLVQDDLG"	Qy	961	GAGCTGTGACGGGTACCAATCGTCTCCATCGAGGACCACTGCAAGGAGATGACTGG	1020
	8826. .12248	Db	339371	GAGCTGTGACGGGTACCAATCGTCTCCATCGAGGACCACTGCAAGGAGATGACTGG	339430
	/gene="pyc"	Qy	1021	GAGGCTTACACCACTTCACCGCAACCATCGGCAAGGTTTCAGATCTGTGGCGACGAC	1080
	/locus_tag="cg0791"	Db	339431	GAGGCTTACACCACTTCACCGCAACCATCGGCAAGGTTTCAGATCTGTGGCGACGAC	339490
	8826. .12248	Qy	1081	TTCTTCTGTCACCAACCTTCGAGCGCTGAAGGAGGATCGCTAAGAGGCTGCCAATCC	1140
	/gene="pyc"	Db	339491	TTCTTCTGTCACCAACCTTCGAGCGCTGAAGGAGGATCGCTAAGAGGCTGCCAATCC	339550
	/locus_tag="cg0791"	Qy	1141	ATCCTGGTTAAGGTGAACAGATCGGTACCCTCACCGAGACCTTCGACGCTGTCGACATG	1200
	/EC_number="6.4.1.1"	Db	339551	ATCCTGGTTAAGGTGAACAGATCGGTACCCTCACCGAGACCTTCGACGCTGTCGACATG	339610
	/codon_start=1	Qy	1201	GCTCACCGCGCAGGCTACACCTCCATGATGTCCACCGTTCGCTGAGACCGAGGACACC	1260
	/transl_table=11	Db	339611	GCTCACCGCGCAGGCTACACCTCCATGATGTCCACCGTTCGCTGAGACCGAGGACACC	339670
	/product="DIHYDROLIPOAMIDE DEHYDROGENASE"	Qy	1261	ACCATTTCTGACCTTCGAGTTGACCTCAACTGTGTGGCCAGATCAAGACTGTGTCTCCAGCA	1320
	/protein_id="CAF19393.1"	Db	339671	ACCATTTCTGACCTTCGAGTTGACCTCAACTGTGTGGCCAGATCAAGACTGTGTCTCCAGCA	339730
	/db_xref="GI:41324912"	Qy	1321	CGTTCCGACCGCTTCGCAAAAGTACACCGCTTCTCCGATTCGAGCAGCTGTGTGGCGAC	1380
	/translations="MAKRIVIIGGSPAGYEAALAGAKYGAETVVIDVGVGSAVTMD	Db	339731	CGTTCCGACCGCTTCGCAAAAGTACACCGCTTCTCCGATTCGAGCAGCTGTGTGGCGAC	339790
	CVPKSGFIAGTKTDLRADDMLNRLGKHALEIDALNIRVKDLAKAQSEDLGOL	Qy	1381	GCCGGCGTCTACGAGGTTCGCAAGGCTTCCCGCTTTTCAGGCTTAAATAAAGCGCTT	1440
	QRSVDMINGVGFEDYNTKQTHYIKVTHSDSESTVECDLVLVATGATPRLLKAE	Db	339791	GCCGGCGTCTACGAGGTTCGCAAGGCTTCCCGCTTTTCAGGCTTAAATAAAGCGCTT	339850
	PDGERLLTRQVTDIELPHTLLVVGSGTVGAFSAFAGELGVKVTVAASRDRLPHD	Qy	1441	TTTGAAGCGGTTAACTCAAGGTTGCCGGCGCTGTTGCTTACTACTGTTACTGTTGT	1500
	DADAADVLETVAERGVSLEKHARVESVTRTDDGGVCRTADREIYGSHALMTGSI	Db	339851	TTTGAAGCGGTTAACTCAAGGTTGCCGGCGCTGTTGCTTACTACTGTTACTGTTGT	339910
	PNTADLGNIGVLAAPSIIKVDVRSRTNIPGYAAGDCTDLFFPLASVAAAMQGRATM				
	YHALGVSVPRLUKTVATVTRPEIAAAGITHAQVDSGVSARVTVLPLATNPRAM				
	YSLRHGFVKLFCRNSGLLIIGVVVAPTASELILPIAVAVTNRLTVADLADTFAVYPS				
	LSGSITFAARQLVQDDLG"				
	8826. .12248				
	/gene="pyc"				
	/locus_tag="cg0791"				
	8826. .12248				
	/gene="pyc"				
	/locus_tag="cg0791"				
	/EC_number="6.4.1.1"				
	/codon_start=1				
	/transl_table=11				
	/product="PYRUVATE CARBOXYLASE"				
	/protein_id="CAF19394.1"				
	/db_xref="GI:41324913"				
	/translations="MSHTSSTLPAPFKKILVANRGEIIVAFRAALLETGAATVAIYPR				
	EDRGSFHRSPASAVRIGTSPVKAYLDIDEIIGAARKYKADAIYGVGFUSENAQL				
	ARECANRGTFTTGTEVLDLTDGKSRVTAAKKAGLPVLAEPSKNIDEIVKSAEG				
	QTPYIFVYVAGGGGMRVSPDELRLKLAETASREAAAFDGVAVYVERAVINPOH				
	IEVOILGDHTVEKHNVDYDCSLQRRHKVETIAPQHLDPDLDRICADAVKPCRSI				
	GYOAGTVEPLDEKHNVPFIEMNPRLOVEHTVTEVVDLKAOMRLAAGATLXEL				
	GLTQDKIKHGAALQCRITTEDPNNRPTDGIITAYRSPGGAGVRLDGAQLGSEIT				
	AHFDMLVKTGCSDFETAVARAKALAEFTVSGVATNIGFLIRALBEDFTSKRIA				
	TGFIADPHLLQAPPADDEQRLDYLADVTNPKPHGVPRKPVAAIPDKLPNIKLDLPL				
	Query Match 100.0%; Score 1578; DB 1; Length 348475;				
	Best Local Similarity 100.0%; Pred. No. 4, 9e-288;				
	Matches 1578; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
Qy	1 GGCTGGGATATGGTATGTTTCCGCCACTAATTTCAACTGATTCGCTCATCGAAACAAGA	338470			
Db	338411 GGCTGGGATATGGTATGTTTCCGCCACTAATTTCAACTGATTCGCTCATCGAAACAAGA	338470			
Qy	61 TTCTGTCACCAATTTGGGTGTAGACGTGATTTGAAGACATTTGATCAGTGAATTTCTAG	120			
Db	338471 TTCTGTCACCAATTTGGGTGTAGACGTGATTTGAAGACATTTGATCAGTGAATTTCTAG	338530			
Qy	121 TTAGTCTCCCAAGTTGGCATAGGAGGCCACAGTGGCTGAAATCATGCAGCTATTCGCTCGC	180			
Db	338531 TTAGTCTCCCAAGTTGGCATAGGAGGCCACAGTGGCTGAAATCATGCAGCTATTCGCTCGC	338590			
Qy	181 GAAATTTCTGACCTCCCGCGGTAAACCAACCGTTCGAGGACAGGTTTTCCTGGATGACGGT	240			
Db	338591 GAAATTTCTGACCTCCCGCGGTAAACCAACCGTTCGAGGACAGGTTTTCCTGGATGACGGT	338650			
Qy	241 TCCACAGGTGTCGAGGTGTTCCATCCGGCGCATCCACCGCGCTCCACGAGGCTCATGAG	300			
Db	338651 TCCACAGGTGTCGAGGTGTTCCATCCGGCGCATCCACCGCGCTCCACGAGGCTCATGAG	338710			
Qy	301 CTGCGTGACGGTGCAGTTCGCTACCTCGGCAAGGCGCTTTTGAAGGAGGTTTGAAGAGTGC	360			
Db	338711 CTGCGTGACGGTGCAGTTCGCTACCTCGGCAAGGCGCTTTTGAAGGAGGTTTGAAGAGTGC	338770			
Qy	361 AACGAGAAATTCGCGCAGAGCTGCTGCTGGCTAGAGGCTGACGATCAGCGCTCATCGAC	420			

QY 1501 GACTATGATCAGGATTATGGCAAGCAGAAGAACTCATAAAGGCTTGTCTCTCTCT 1560
|||||
Db 339911 GACTATGATCAGGATTATGGCAAGCAGAAGAACTCATAAAGGCTTGTCTCTCTCT 339970
|||||
QY 1561 CAAGCAGGAAACGTGCTT 1578
|||||
Db 339971 CAAGCAGGAAACGTGCTT 339988
|||||
RESULT 6
AX127146 AX127146 349980 bp DNA linear (PAT 14 MAY 2001)
LOCUS AX127146 Sequence 7062 from Patent EP1108790.
DEFINITION AX127146 AX114121
ACCESSION AX127146.1 GI:14041134
VERSION
KEYWORDS
SOURCE Corynebacterium glutamicum
ORGANISM Corynebacterium glutamicum
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Corynebacteriaceae; Corynebacterium.
REFERENCE 1
AUTHORS Nakagawa,S., Mizoguchi,H., Ando,S., Hayashi,M., Ochiai,K.,
Yokoi,H., Tateishi,N., Senoh,A., Ikeda,M. and Ozaki,A.
TITLE Novel polynucleotides
JOURNAL Patent: EP 1108790-A 7062 20-JUN-2001;
KYOWA HAKKO KOGYO CO., LTD. (JP)
FEATURES
source Location/Qualifiers
1. 349980
/organism="Corynebacterium glutamicum"
/mol_type="genomic DNA"
/db_xref="taxon:1718"
/note="Seq 1 to long (3.309.400) split in 11, seq 7062
0.900.001 1.249.980"
ORIGIN
Query Match 100.0%; Score 1578; DB 6; Length 349980;
Best Local Similarity 100.0%; Pred. No. 4.9e-288;
Matches 1578; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGCTGGGGATATGGGTAGTTTTCGCCCACTAATTTCAACTGATGTCCTCATCGAAACAAGA 60
Db 134799 GGCTGGGGATATGGGTAGTTTTCGCCCACTAATTTCAACTGATGTCCTCATCGAAACAAGA 134858
|||||
QY 61 TTCTGTCAACAATTTGGGTAGAGCTGATTGAAGACATTTTGATCAGCTGAATATTTCTAG 120
Db 134859 TTCTGTCAACAATTTGGGTAGAGCTGATTGAAGACATTTTGATCAGCTGAATATTTCTAG 134918
|||||
QY 121 TTAGCTCCCAAGTTGGCATAGGAGGCCACAGTGGCTGAAATCATGCACTATTTCGCTCGC 180
Db 134919 TTAGCTCCCAAGTTGGCATAGGAGGCCACAGTGGCTGAAATCATGCACTATTTCGCTCGC 134978
|||||
QY 181 GAAATTCGACTCCCGGGTAACCCACCGTCAGGAGAGGTTTTCCTGGATGACGGT 240
Db 134979 GAAATTCGACTCCCGGGTAACCCACCGTCAGGAGAGGTTTTCCTGGATGACGGT 135038
|||||
QY 241 TCCACAGCTGTGCGAGGTGTTCCATCCGGGCGATCCACCGCGTCCACGAGGCTCATGAG 300
Db 135039 TCCACAGCTGTGCGAGGTGTTCCATCCGGGCGATCCACCGCGTCCACGAGGCTCATGAG 135098
|||||
QY 301 CTGCGTGACGGTGGCGATCGCTACTCTGGGCAAGGGCGTTTGAAGGAGGTTGAAGAGCGTC 360
Db 135099 CTGCGTGACGGTGGCGATCGCTACTCTGGGCAAGGGCGTTTGAAGGAGGTTGAAGAGCGTC 135158
|||||
QY 361 AACGAAGAAATCGCGGACGAGCTCGCTGGCTAGAGGCTGACATCAGCGCTCATCGAC 420
Db 135159 AACGAAGAAATCGCGGACGAGCTCGCTGGCTAGAGGCTGACATCAGCGCTCATCGAC 135218
|||||
QY 421 GAAGCAATGATCAAGCTTGATGGCAAGCGCAAGTCCCGCTGGGTGCAAGCGCAATC 480
Db 135219 GAAGCAATGATCAAGCTTGATGGCAAGCGCAAGTCCCGCTGGGTGCAAGCGCAATC 135278
|||||
QY 481 CTTGGTGTTCATGGCTGTTGCAAGAGCTGCTGATTCGGCAGGCGCTCCCACTGCTTC 540
|||||

Db 135279 CTTGGTGTTCATGGCTGTTGCAAAAGGCTGCTGCTGATTCGCGAGGCTCCCACTGTTC 135338
|||||
QY 541 CGCTACATCGGTGGACCAACAGCACAGCTTCTCCAGTCCCAATGATGAACATCATCAAC 600
|||||
Db 135339 CGCTACATCGGTGGACCAACAGCACAGCTTCTCCAGTCCCAATGATGAACATCATCAAC 135398
|||||
QY 601 GGTGGGGCTCACGCTGACTCCGGTGTGAGCTTCAGGAATTCATGATCGCTCCAAATCGGT 660
|||||
Db 135399 GGTGGGGCTCACGCTGACTCCGGTGTGAGCTTCAGGAATTCATGATCGCTCCAAATCGGT 135458
|||||
QY 661 GCAGAGACTTCTCTGAGGCTCTCCGCAACGGCGCGAGGTCTTACACGCACTGAAAGTCC 720
|||||
Db 135459 GCAGAGACTTCTCTGAGGCTCTCCGCAACGGCGCGAGGTCTTACACGCACTGAAAGTCC 135518
|||||
QY 721 GTCATCAAGGAAAGGCGCTGTCACCGGACTTTGGCGATGAGGGCGGCTTCGCTCCTTCC 780
|||||
Db 135519 GTCATCAAGGAAAGGCGCTGTCACCGGACTTTGGCGATGAGGGCGGCTTCGCTCCTTCC 135578
|||||
QY 781 GTCGGCTCCACCGGTGAGGCTCTTGACCTTATCGTTGAGGCAATTCGAGAAGGCTGGCTTC 840
|||||
Db 135579 GTCGGCTCCACCGGTGAGGCTCTTGACCTTATCGTTGAGGCAATTCGAGAAGGCTGGCTTC 135638
|||||
QY 841 ACCCGAGGCAAGGACATCGCTCTTGCTCTGAGAGTTCCTCTGAGTTCCTCAAGGAC 900
|||||
Db 135639 ACCCGAGGCAAGGACATCGCTCTTGCTCTGAGAGTTCCTCTGAGTTCCTCAAGGAC 135698
|||||
QY 901 GGCACCTACCACCTTCGAAGGTGGCGAGCACTCCGCGAGCTGAGATGGCAAGCTTACGCT 960
|||||
Db 135699 GGCACCTACCACCTTCGAAGGTGGCGAGCACTCCGCGAGCTGAGATGGCAAGCTTACGCT 135758
|||||
QY 961 GAGCTCGTTGACCGGTACCCCAATCGTCTCCATCGAGACCCACTGCAAGGAGAGTACTGG 1020
|||||
Db 135759 GAGCTCGTTGACCGGTACCCCAATCGTCTCCATCGAGACCCACTGCAAGGAGAGTACTGG 135818
|||||
QY 1021 GAGGTTTACCAACCTCACCGCAACCATCGGGGACAAGGTTTCAAGTTCGTTGGCGACGAC 1080
|||||
Db 135819 GAGGTTTACCAACCTCACCGCAACCATCGGGGACAAGGTTTCAAGTTCGTTGGCGACGAC 135878
|||||
QY 1081 TTCTTGTTCACCAACCTCGAGCGCTGAAGGAGGCAATCGCTAAGAAGGCTGCCCACTCC 1140
|||||
Db 135879 TTCTTGTTCACCAACCTCGAGCGCTGAAGGAGGCAATCGCTAAGAAGGCTGCCCACTCC 135938
|||||
QY 1141 ATCTGTGTTAAGGTGAACACAGATCGGTACCTTCACCGAGACCTTCGACGCTGTCGACATG 1200
|||||
Db 135939 ATCTGTGTTAAGGTGAACACAGATCGGTACCTTCACCGAGACCTTCGACGCTGTCGACATG 135998
|||||
QY 1201 GCTCACCGCAGGCTACACCTCCATGATGTCCACCGTTCGGGTGAGACCGAGGACACC 1260
|||||
Db 135999 GCTCACCGCAGGCTACACCTCCATGATGTCCACCGTTCGGGTGAGACCGAGGACACC 136058
|||||
QY 1261 ACCATTGCTGACCTCGCAGTTGCACTCAACTGTGGCCAGATCAAGACTTGGTGTCTCCAGCA 1320
|||||
Db 136059 ACCATTGCTGACCTCGCAGTTGCACTCAACTGTGGCCAGATCAAGACTTGGTGTCTCCAGCA 136118
|||||
QY 1321 CGTTCCAGCCTGTGCGCAAAAGTACAACAGCTTCTCCGCACTCGAGAGCTGCTTTGGCGAC 1380
|||||
Db 136119 CGTTCCAGCCTGTGCGCAAAAGTACAACAGCTTCTCCGCACTCGAGAGCTGCTTTGGCGAC 136178
|||||
QY 1381 GCCGGGCTTACGAGGTCGAGCGCATTCGACGCTTTCAGGGCTAAATAAAGCGCTT 1440
|||||
Db 136179 GCCGGGCTTACGAGGTCGAGCGCATTCGACGCTTTCAGGGCTAAATAAAGCGCTT 136238
|||||
QY 1441 TTGACGCCCGGTAACTCAAGGTTGCGGGCGCTGCTGCTTACTACTGTTACTGCTGT 1500
|||||
Db 136239 TTGACGCCCGGTAACTCAAGGTTGCGGGCGCTGCTGCTTACTACTGTTACTGCTGT 136298
|||||
QY 1501 GACTATGATCAGGATTATGGCAAGCAGAAGAACTCATAAAGGCTTGTCTCTCTCTCT 1560
|||||
Db 136299 GACTATGATCAGGATTATGGCAAGCAGAAGAACTCATAAAGGCTTGTCTCTCTCTCT 136358
|||||
QY 1561 CAAGCAGGAAACGTGCTT 1578
|||||

Dn	136359	CAAGCAGGGAACGTGCTT	136376
RESULT 7			
AR490886	LOCUS	AR490886	Sequence 3 from patent US 6713289.
DEFINITION	Accession	AR490886	Version
KEYWORDS	Keywords	Unknown.	GI:47258405
SOURCE	Source	Unknown.	Organism
REFERENCE	Reference	Unclassified.	Authors
Möckel,B., Pfeifferle,W., Hermann,T., Puhler,A., Kalinowski,J. and Bathe,B.	Nucleotide sequences which code for the eno gene	Patent: US 6713289-A 30-MAR-2004;	Location/Qualifiers
FEATURES	Feature	1..1578	/organism="unknown"
ORIGIN	Origin	/mol_type="genomic DNA"	
Query Match	Score	99.9%;	DB 6; Length 1578;
Best Local Similarity	Pred.No.	99.9%;	1.4e-287;
Matches 157;	Conservative	0;	Mismatches 1; Indels 0; Gaps 0;
Qy	1	GGCTGGGGATATCGGTAGTTTTCGCCACTAAATTTCACCTGATTTGCCTCATCGAACAAGA	60
Dn	1	GGCTGGGGATATCGGTAGTTTTCGCCACTAAATTTCACCTGATTTGCCTCATCGAACAAGA	60
Qy	61	TTCCGTCACAAATGGGTGATAGCGTGATTAAGACATTTGATCAGTGAATAATCTTAG	120
Dn	61	TTCCGTCACAAATGGGTGATAGCGTGATTAAGACATTTGATCAGTGAATAATCTTAG	120
Qy	121	TTAGCTCCCCAAGTTGGCATAGAGGCGCACAGTGGTGAATCATCGACGTATTCGCTCGC	180
Dn	121	TTAGCTCCCCAAGTTGGCATAGAGGCGCACAGTGGTGAATCATCGACGTATTCGCTCGC	180
Qy	181	GAAATTCGACTCCCGCGGTAAACCACCGTGGAGGAGAGGTTTTCTGGATGACGGT	240
Dn	181	GAAATTCGACTCCCGCGGTAAACCACCGTGGAGGAGAGGTTTTCTGGATGACGGT	240
Qy	241	TCCCACGGTGTCCAGGTGTTTCCATCCGGCGCATCCACGGCGGTCCACGAGGCTCATGAG	300
Dn	241	TCCCACGGTGTCCAGGTGTTTCCATCCGGCGCATCCACGGCGGTCCACGAGGCTCATGAG	300
Qy	301	CTCGGTGA CGGTGGGATCGCTACCTGGGCAAGGGCGTTTTGAAGGCAAGTTGAAAACGTC	360
Dn	301	CTCGGTGA CGGTGGGATCGCTACCTGGGCAAGGGCGTTTTGAAGGCAAGTTGAAAACGTC	360
Qy	361	AAGCAAGAAATCGGACACGAGCTCGCTGCGCTAGAGGCTGACGATCAGCGCTCATCGAC	420
Dn	361	AAGCAAGAAATCGGACACGAGCTCGCTGCGCTAGAGGCTGACGATCAGCGCTCATCGAC	420
Qy	421	GAAGCAATGATCAAGCTTGAATGACACCGCAACAAAGTCCCGCTGGGTGCAAAACGCAATC	480
Dn	421	GAAGCAATGATCAAGCTTGAATGACACCGCAACAAAGTCCCGCTGGGTGCAAAACGCAATC	480
Qy	481	CTTGGTGTTCATGCTGCTTGAAGGCTGCTGATTCGCGAGGCTCCCACTGTTTC	540
Dn	481	CTTGGTGTTCATGCTGCTTGAAGGCTGCTGATTCGCGAGGCTCCCACTGTTTC	540
Qy	541	CGCTACATCGGTGGACCAAACGACACGCTTCTTCCAGTTTCCAATGATGAACATCATCAAC	600
Dn	541	CGCTACATCGGTGGACCAAACGACACGCTTCTTCCAGTTTCCAATGATGAACATCATCAAC	600
Qy	601	GGTGGGCTCAGCTGACCTCCGGTGTGACGTTTCAAGATTCATGATCGCTCCCAATCGGT	660
Dn	601	GGTGGGCTCAGCTGACCTCCGGTGTGACGTTTCAAGATTCATGATCGCTCCCAATCGGT	660
Qy	661	GCAGACCTTCTCTGAGGCTCTCCGCAACGGCGGAGGTTACACGACCTGAAGTCC	720

Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Corynebacteriaceae; Corynebacterium.

1
Zelder, O., Pompejus, M., Schroeder, H., Kroeger, B., Klopprogge, C. and
Haberkauer, G.
Genes encoding for carbon metabolism and energy-producing proteins
Patent: WO 03040291-A 7 15-MAY-2003;
BASF AKTIENGESSELLSCHAFT (DE)
Location/Qualifiers

source
1. 1405
/organism="Corynebacterium glutamicum"
/mol_type="unassigned DNA"
/db_xref="taxon:1718"
101. .1378
/note="unnamed protein product; RXA00235"
/codon_start=1
/transl_table=1
/protein_id="CAD99859.1"
/db_xref="GI:32257759"
/translation="MAEIMHVFAREIILDSRGNPTVEAEVFLDDSGHGVAGVPSGASTG
VHEAHLRDGDRYLGKVLKAVENVEIGDELAGEADDORLIDAMIKLDGTANK
SRLGANAILGVSMVAKAADSAGLPLFRVIGGPNHVLVPMWNIIITGGHADSGVD
VOEPMIPIAGETSEALRNGAEYHAKSVIKKGLSTGLDGGFPAPSVGSTREAL
DLVLEAKAGFTPGKIDIALADLVASSEFFKDYHPEGGHSAENVAELVDAY
PIVSIEDPLEQEDMEGYTNLTATIGDKVQIVGDDFFVTNPERLKEGTAKKAANSILVK
VNQIGTUTETFDVMAHRAGYTSMMSHRSGETEDTITADLVALNCGIKTGPAPRS
DRVAKYNQLRIEQLLDAGVYAGRSAPPRFG"

ORIGIN

Query Match 88.9%; Score 1403.4; DB 6; Length 1405;
Best Local Similarity 99.9%; Pred. No. 6.6e-255;
Matches 1404; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy	51	CGAACAAAGATTCGTGCAACAATTTGGGTGTAGAGCTGATTCAAGACATTTTGCACGTGA	110
Db	1	CGAACAAAGATTCGTGCAACAATTTGGGTGTAGAGCTGATTCAAGACATTTTGCACGTGA	60
Qy	111	ATAATTCCTAGTTAGCTCCCAAGTTGGCATAGAGGCCACAGTGGCTGAATCATGCACGT	170
Db	61	ATAATTCCTAGTTAGCTCCCAAGTTGGCATAGAGGCCACAGTGGCTGAATCATGCACGT	120
Qy	171	ATTCGCTCGGAAATTTCTCGACTCCCGCGTAAACCCAAACCGTCGAGGCGAGGTTTTCCT	230
Db	121	ATTCGCTCGGAAATTTCTCGACTCCCGCGTAAACCCAAACCGTCGAGGCGAGGTTTTCCT	180
Qy	231	GGATGACGGTTCACAGGTTCGACAGTGTTCATTCGGCGCATCCACCGCGTCCACGA	290
Db	181	GGATGACGGTTCACAGGTTCGACAGTGTTCATTCGGCGCATCCACCGCGTCCACGA	240
Qy	291	GGCTCATGAGCTGGTGAACGGTGGCATCGCTACTCGGCAAGGCGCTTTTGAAGGCGAT	350
Db	241	GGCTCATGAGCTGGTGAACGGTGGCATCGCTACTCGGCAAGGCGCTTTTGAAGGCGAT	300
Qy	351	TGAAAACGTCAACGAAGAAATCGGCGACGAGCTCGCTCGCTAGAGGCTGACGATCAGCG	410
Db	301	TGAAAACGTCAACGAAGAAATCGGCGACGAGCTCGCTCGCTAGAGGCTGACGATCAGCG	360
Qy	411	CCTCATCGAAGCAATGATCAAGCTTGATGGCACCGCCAAAGTCCCGCTGGGTGC	470
Db	361	CCTCATCGAAGCAATGATCAAGCTTGATGGCACCGCCAAAGTCCCGCTGGGTGC	420
Qy	471	AAACGCAATCCTCGTGTTCATGGCTGTTTGCAGAGGCTGCTGATTCGCGAGGCGCT	530
Db	421	AAACGCAATCCTCGTGTTCATGGCTGTTTGCAGAGGCTGCTGATTCGCGAGGCGCT	480
Qy	531	CCCACTGTTCGCTACATCGGTGGACCAACGACACGCTTCTCCAGTTCCAATGATGAA	590
Db	481	CCCACTGTTCGCTACATCGGTGGACCAACGACACGCTTCTCCAGTTCCAATGATGAA	540
Qy	591	CATCATCAACGGTGGCGCTCAGCTGACTCCGGTGTTCAGCTTCAGGAATTCATGATCGC	650
Db	541	CATCATCAGCGTGGCGCTCAGCTGACTCCGGTGTTCAGCTTCAGGAATTCATGATCGC	600

651 TCCAATCGTGCAGAGACCTTCTCTGAGGCTCTCCGCAACGGCGCGGAGGTCTACACGC

710 TCCAATCGTGCAGAGACCTTCTCTGAGGCTCTCCGCAACGGCGCGGAGGTCTACACGC

770 ACTGAAGTCCGTCTATCAAGGAAAGGCGCTGTCCACCGGACTTGGCGATGAGGCGGCTT

830 ACTGAAGTCCGTCTATCAAGGAAAGGCGCTGTCCACCGGACTTGGCGATGAGGCGGCTT

890 CGCTCTCTCGTCCGCTCCACCGGCTTCTTGAACCTTATCGTTGAGGCAATCGAGAA

950 CGCTCTCTCGTCCGCTCCACCGGCTTCTTGAACCTTATCGTTGAGGCAATCGAGAA

1010 CGCTCTCTCGTCCGCTCCACCGGCTTCTTGAACCTTATCGTTGAGGCAATCGAGAA

1070 AGATGACTGGAGGGTTTACACCACTCACCGCAACCATCGCGCAACGAGTTTCAGATCGT

1130 AGATGACTGGAGGGTTTACACCACTCACCGCAACCATCGCGCAACGAGTTTCAGATCGT

1190 TGCGGACGACTTCTTCTGTCACCAACCTTGAGCGCTGAAAGGAGGCGATCGCTAAGAGGC

1250 TGCGGACGACTTCTTCTGTCACCAACCTTGAGCGCTGAAAGGAGGCGATCGCTAAGAGGC

1310 TGCGGACGACTTCTTCTGTCACCAACCTTGAGCGCTGAAAGGAGGCGATCGCTAAGAGGC

1370 TGCGGACGACTTCTTCTGTCACCAACCTTGAGCGCTGAAAGGAGGCGATCGCTAAGAGGC

1430 TGCGGACGACTTCTTCTGTCACCAACCTTGAGCGCTGAAAGGAGGCGATCGCTAAGAGGC

1490 TGCGGACGACTTCTTCTGTCACCAACCTTGAGCGCTGAAAGGAGGCGATCGCTAAGAGGC

1550 TGCGGACGACTTCTTCTGTCACCAACCTTGAGCGCTGAAAGGAGGCGATCGCTAAGAGGC

1610 TGCGGACGACTTCTTCTGTCACCAACCTTGAGCGCTGAAAGGAGGCGATCGCTAAGAGGC

1670 TGCGGACGACTTCTTCTGTCACCAACCTTGAGCGCTGAAAGGAGGCGATCGCTAAGAGGC

1730 TGCGGACGACTTCTTCTGTCACCAACCTTGAGCGCTGAAAGGAGGCGATCGCTAAGAGGC

1790 TGCGGACGACTTCTTCTGTCACCAACCTTGAGCGCTGAAAGGAGGCGATCGCTAAGAGGC

1850 TGCGGACGACTTCTTCTGTCACCAACCTTGAGCGCTGAAAGGAGGCGATCGCTAAGAGGC

1910 TGCGGACGACTTCTTCTGTCACCAACCTTGAGCGCTGAAAGGAGGCGATCGCTAAGAGGC

1970 TGCGGACGACTTCTTCTGTCACCAACCTTGAGCGCTGAAAGGAGGCGATCGCTAAGAGGC

2030 TGCGGACGACTTCTTCTGTCACCAACCTTGAGCGCTGAAAGGAGGCGATCGCTAAGAGGC

2090 TGCGGACGACTTCTTCTGTCACCAACCTTGAGCGCTGAAAGGAGGCGATCGCTAAGAGGC

2150 TGCGGACGACTTCTTCTGTCACCAACCTTGAGCGCTGAAAGGAGGCGATCGCTAAGAGGC

2210 TGCGGACGACTTCTTCTGTCACCAACCTTGAGCGCTGAAAGGAGGCGATCGCTAAGAGGC

2270 TGCGGACGACTTCTTCTGTCACCAACCTTGAGCGCTGAAAGGAGGCGATCGCTAAGAGGC

2330 TGCGGACGACTTCTTCTGTCACCAACCTTGAGCGCTGAAAGGAGGCGATCGCTAAGAGGC

2390 TGCGGACGACTTCTTCTGTCACCAACCTTGAGCGCTGAAAGGAGGCGATCGCTAAGAGGC

2450 TGCGGACGACTTCTTCTGTCACCAACCTTGAGCGCTGAAAGGAGGCGATCGCTAAGAGGC

RESULT 9
AX064945
LOCUS
DEFINITION
Sequence 71 from Patent WO0100844.
AX064945
ACCESSION
AX064945.1
GI:12542657

KEYWORDS
SOURCE
ORGANISM

Corynebacterium glutamicum
Corynebacterium glutamicum
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Corynebacteriaceae; Corynebacterium.

REFERENCE
AUTHORS
TITLE

1
Pompejus, M., Kroeger, B., Schroeder, H., Zelder, O. and Haberkauer, G.
Corynebacterium glutamicum genes encoding proteins involved in
carbon metabolism and energy production
Patent: WO 0100844-A 71 04-JAN-2001;
JOURNAL

FEATURES		BASF AKTIENSELLSCHAFT (DE)
Source	Location/Qualifiers	1. .1398
	/organism="Corynebacterium glutamicum"	
	/mol_type="unassigned DNA"	/db_xref="taxon:1718"
	101..1378	
CDS	/notes="unnamed protein product; RXA00235"	
	/codon_start=1	
	/transl_table=11	
	/protein_id="CAC25712.1"	
	/db_xref="GI:12542658"	
	/translation="MARIMHFAREILDSGNPTVAEVEFLDDSGHVGAVPSGASTG VHAHEURDGGRLGLGVLUKAVENNEEIGDELAGLEADDQRLIDEMIKLDTANK SRIGANAILGVSAVAKAASAGYLPKFRVYGGPNNAHLVPMNNIITGAHDSGVD VQFMFAPIGAEFTSEALRNGAEVYLFKRIKELGSTGLGDEGDFAPSVGSTRAL DLIVEATEKAGFTPGKIDIALDVAISSEFPKDGTYHPEGQHSAAEMANYAELVDAY PIVISEDPLEQDWEVNTLATIGDKVQIGDDFFVTNPERLKEGIKAKAANSILVK VNOIGTLTETFDVADMAHRAGYTSMHSRSETEDTTIADLAVALNCGKQKIGAPARS DRVAKYNQLRIEQLLDGAGVYAGRSAPPRQG"	
	ORIGIN	
	Query Match	88.5%; Score 1396.4; DB 6; Length 1398;
	Best Local Similarity	99.9%; Pred. No. 1.4e-253;
	Matches 1397; Conservative	0; Mismatches 1; Indels 0; Gaps 0;
Qy	51	CGAAACAAGATTCTCGTGAACAATTGGGTGTAGACGTGATTGAAGACATTTGATCACGTGA 110
Db	1	CGAAACAAGATTCTCGTGAACAATTGGGTGTAGACGTGATTGAAGACATTTGATCACGTGA 60
Qy	111	ATAATTCTAGTTAGTCCCAAGTTGGCATAGGAGCCACAGTGGCTGAATCATGACAGT 170
Db	61	ATAATTCTAGTTAGTCCCAAGTTGGCATAGGAGCCACAGTGGCTGAATCATGACAGT 120
Qy	171	ATTGGCTCGCAAAATTTCTCGACTCCCGCGGTAAACCCGTCGAGGCGAGGTTTTCT 230
Db	121	ATTGGCTCGCAAAATTTCTCGACTCCCGCGGTAAACCCGTCGAGGCGAGGTTTTCT 180
Qy	231	GGATGACGGTTCCACCGGTGTCGACGTTGTCATCCGGCGCATCCACCGCGTCCACGA 290
Db	181	GGATGACGGTTCCACCGGTGTCGACGTTGTCATCCGGCGCATCCACCGCGTCCACGA 240
Qy	291	GGCTCATGAGTGGGTGCGGTGGCGATCGCTACTCGGCGAAGCGGTTTTGAAGGCGAGT 350
Db	241	GGCTCATGAGTGGGTGCGGTGGCGATCGCTACTCGGCGAAGCGGTTTTGAAGGCGAGT 300
Qy	351	TGAAAACGTCAACGAAGAAATCGGCGACGAGCTCGTGGCTTAGAGGCTGACGATCAGCG 410
Db	301	TGAAAACGTCAACGAAGAAATCGGCGACGAGCTCGTGGCTTAGAGGCTGACGATCAGCG 360
Qy	411	CCTCATCGACGAAGCAATGATCAAGCTTGATGGCAACCGCCCAACAGTCCCGCTGGGTGC 470
Db	361	CCTCATCGACGAAGCAATGATCAAGCTTGATGGCAACCGCCCAACAGTCCCGCTGGGTGC 420
Qy	471	AAACCAATCTTGTGTTTTCATCGCTGTTGCAAAAGGCTGCTGCTGATTCCGAGGCGCT 530
Db	421	AAACCAATCTTGTGTTTTCATCGCTGTTGCAAAAGGCTGCTGCTGATTCCGAGGCGCT 480
Qy	531	CCCACTGTCGCTACATCGGTGGACCAAAACGACACGTTCTTCAGTTCCAAATGATGAA 590
Db	481	CCCACTGTCGCTACATCGGTGGACCAAAACGACACGTTCTTCAGTTCCAAATGATGAA 540
Qy	591	CATCATCAACGGTGGGCTCAGCTGACTCCGGTGTTCAGTTTCAGGAAATCATGATCGC 650
Db	541	CATCATCAACGGTGGGCTCAGCTGACTCCGGTGTTCAGTTTCAGGAAATCATGATCGC 600
Qy	651	TCCAATCGGTGACAGACCTTCTCTGAGGCTCTCCGCAACCGCGCGGAGGTTCTACACGC 710
Db	601	TCCAATCGGTGACAGACCTTCTCTGAGGCTCTCCGCAACCGCGCGGAGGTTCTACACGC 660
Qy	711	ACTGAAGTCCGTCATCAAGGAAAAGGCTGTGTCACCGGACTTGGCGATGAGGGCGGTT 770
Db	661	ACTGAAGTCCGTCATCAAGGAAAAGGCTGTGTCACCGGACTTGGCGATGAGGGCGGTT 720
RESULTS		RESULT 10
BD163286	LOCUS	BD163286
	DEFINITION	Novel polynucleotide.
	ACCESSION	BD163286
	VERSION	BD163286.1 GI:27869050
	KEYWORDS	JP 2002191370-A/1085.
SOURCE	unidentified	
	ORGANISM	unclassified.
REFERENCE	1 (bases 1 to 1275)	
	AUTHORS	Nakagawa, S., Mizoguchi, H., Ando, S., Hayashi, M., Ochiai, K., Yokoi, H., Tateishi, N., Senoo, A., Ikeda, M. and Ozaki, A.
TITLE	Novel polynucleotide	
	JOURNAL	Patent: JP 2002191370-A 1085 09-JUL-2002;
COMMENT	OS	KYOWA HAKKO KOGYO CO LTD
	PN	Corynebacterium glutamicum
	PD	JP 2002191370-A/1085
	PF	09-JUL-2002
	PI	15-DEC-2000 JP 2000405096
	PI	SATOSHI NAKAGAWA, HIROSHI MIZOGUCHI, SEIKO ANDO, MIKIO HAYASHI, KEIKO OCHIAI,
	PI	HARUHIKO YOKOI, NAKO TATEISHI, AKIHIRO SENOO, MASATO IKEDA, AKIO

QY	211	GTCCAGGCGAGAGGTTTTCTTGGATGACCGGTTCCACGGTGTCCGAGGTGTTCATCCGGC	270
DB	61	GTCCAGGCGAGAGGTTTTCTTGGATGACCGGTTCCACGGTGTTCATCCGGC	120
QY	271	GCATCCACCGCGTCCACGAGGCTCATGAGTCGGTGA CGGTGGCGATCGTACCTGGGC	330
DB	121	GCATCCACCGCGTCCACGAGGCTCATGAGTCGGTGA CGGTGGCGATCGTACCTGGGC	180
QY	331	AAGGGCGTTTTGAAGGCAGTTGAAAACGTCAA CGAGAAATCGGCGACGAGCTCGTGGC	390
DB	181	AAGGGCGTTTTGAAGGCAGTTGAAAACGTCAA CGAGAAATCGGCGACGAGCTCGTGGC	240
QY	391	CTAGAGGCTGACGATCAGCGCTCATCGACGAAGCAATGATCAAGCTTGATGCACCGCC	450
DB	241	CTAGAGGCTGACGATCAGCGCTCATCGACGAAGCAATGATCAAGCTTGATGCACCGCC	300
QY	451	AACAAGTCCCGCTGGGTGCAACGCAATCTTTGGTGTTCATATGGCTGTGTCAAAAGGT	510
DB	301	AACAAGTCCCGCTGGGTGCAACGCAATCTTTGGTGTTCATATGGCTGTGTCAAAAGGT	360
QY	511	GCTGCTGATTCGCGACGCTCCACTGTTCCGCTACATCGGTGGACCAACGACACAGTT	570
DB	361	GCTGCTGATTCGCGACGCTCCACTGTTCCGCTACATCGGTGGACCAACGACACAGTT	420
QY	571	CTTCCAGTTCCAATGATGAACATCATCAACGGTGGCGCTCACGCTGACTCCGGTGTTCAC	630
DB	421	CTTCCAGTTCCAATGATGAACATCATCAACGGTGGCGCTCACGCTGACTCCGGTGTTCAC	480
QY	631	GTTCCAGGAATTCATGATCGCTCCAAATCGGTGCAGAGACCTTCTCTGAGGCTCTCCGCAAC	690
DB	481	GTTCCAGGAATTCATGATCGCTCCAAATCGGTGCAGAGACCTTCTCTGAGGCTCTCCGCAAC	540
QY	691	GGCGCGAGAGTCTACACGCACTGAAAGTCGCTCATCAAGGAAAGGGCCCTGTCCACCGGA	750
DB	541	GGCGCGAGAGTCTACACGCACTGAAAGTCGCTCATCAAGGAAAGGGCCCTGTCCACCGGA	600
QY	751	CTTGGCGATGAGGGCGGCTTCGCTCCTTCGTCGGCTCCACCGTGAAGGCTCTTGACCTT	810
DB	601	CTTGGCGATGAGGGCGGCTTCGCTCCTTCGTCGGCTCCACCGTGAAGGCTCTTGACCTT	660
QY	811	ATCGTTGAGGCAATCGAAGAGGCTGGCTTCACCCAGGCGAAGACATCGCTCTTGCTCTG	870
DB	661	ATCGTTGAGGCAATCGAAGAGGCTGGCTTCACCCAGGCGAAGACATCGCTCTTGCTCTG	720
QY	871	GACGTTGCTTCTCTGAGTTCTTCAAGGACGGACCTTACCACTTCGAAGGTGGCAGCAC	930
DB	721	GACGTTGCTTCTCTGAGTTCTTCAAGGACGGACCTTACCACTTCGAAGGTGGCAGCAC	780
QY	931	TCCGACGTGAGTGGCAACGTTTTACGCTGAGCTCGTTGACGGTACCCAAATGCTCTCC	990
DB	781	TCCGACGTGAGTGGCAACGTTTTACGCTGAGCTCGTTGACGGTACCCAAATGCTCTCC	840
QY	991	ATCAGGAGCCACTGCAGGAAGATGACTGGGAGGTTACACCAACCTCACCGCAACCATC	1050
DB	841	ATCAGGAGCCACTGCAGGAAGATGACTGGGAGGTTACACCAACCTCACCGCAACCATC	900
QY	1051	GGCGACAAGGTTAGATCGTTGGCGACGACTCTTCGTTCAACCAACCTGAGCGGCTGAAG	1110
DB	901	GGCGACAAGGTTAGATCGTTGGCGACGACTCTTCGTTCAACCAACCTGAGCGGCTGAAG	960
QY	1111	GAGGGCATCGTAAAGGCTGCCAATCCATCTCGTTAAGGTGAACAGATCGGTACC	1170
DB	961	GAGGGCATCGTAAAGGCTGCCAATCCATCTCGTTAAGGTGAACAGATCGGTACC	1020
QY	1171	CTCACCGAGACTTCGAGCTGTGGACATGGCTCACCGCGAGGCTACACTCCATGATG	1230
DB	1021	CTCACCGAGACTTCGAGCTGTGGACATGGCTCACCGCGAGGCTACACTCCATGATG	1080
QY	1231	TCCCAACGTTCCGGTGAACCGAGGACACCAATTCGTTGCTGACTCGAGTTGCATCAAC	1290
DB	1081	TCCCAACGTTCCGGTGAACCGAGGACACCAATTCGTTGCTGACTCGAGTTGCATCAAC	1140

Qy	1291	TGTGGCCAGATCAAGACTGGTGTCTCCAGCACGTTTCGACCGTGTCCGAAAGTACAACCCAG	1355
Db	1141	TGTGGCCAGATCAAGACTGGTGTCTCCAGCACGTTTCGACCGTGTCCGAAAGTACAACCCAG	1200
Qy	1351	CTTCTCCGCATCAGCAGCTGCTTGGCGACGCGCGGTCTACGCAGGTTCGACGCGCATTC	1410
Db	1201	CTTCTCCGCATCAGCAGCTGCTTGGCGACGCGCGGTCTACGCAGGTTCGACGCGCATTC	1260
Qy	1411	CCACGCTTTTCAGGCG	1425
Db	1261	CCACGCTTTTCAGGCG	1275
RESULT 12			
AP005217			
LOCUS	AP005217	300750 bp	DNA linear BCT 24-JUL-2003
DEFINITION	Corynebacterium efficiens YS-314 DNA, complete genome, section 4/11.		
ACCESSION	AP005217	BA000035	
VERSION	AP005217.1	GI:23492722	
KEYWORDS	Corynebacterium efficiens YS-314		
SOURCE	Corynebacterium efficiens YS-314		
ORGANISM	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacterineae; Corynebacteriaceae; Corynebacterium.		
REFERENCE	1	Nishio Y., Nakamura, Y., Kawarabayasi, Y., Usuda, Y., Kimura, E., Sugimoto, S., Matsui, K., Yamagishi, A., Kikuchi, H., Ikeo, K. and Gojobori, T.	
AUTHORS	Comparative complete genome sequence analysis of the amino acid replacements responsible for the thermostability of Corynebacterium efficiens		
TITLE	Genome Res. 13 (7), 1572-1579 (2003)		
JOURNAL	22723752		
MEDLINE	12840036		
PUBMED	2 (bases 1 to 300750)		
REFERENCE	Kawarabayasi, Y., Yamazaki, J., Hino, Y., Kikuchi, H. and Director-General of Biotechnology Center.		
AUTHORS	Submitted (17-MAY-2002) Director-General of Biotechnology Center, National Institute of Technology and Evaluation, Biotechnology Center; Nishihara 2-49-10, Shibuya-ku, Tokyo 151-0066, Japan (E-mail:biocente.go.jp, Tel:81-3-3481-1933, Fax:81-3-3481-8424)		
TITLE	Kawarabayasi, Y. is officially affiliated with the National Institute of Advanced Industrial Science and Technology, Tsukuba, Ibaraki, 305-8566 Japan		
JOURNAL	Nakamura, Y., Ikeo, K., Suzuki, M. and Mashima, J. are at the National Institute of Genetics, Mishima, Shizuoka, 411-8540 Japan		
COMMENT	Icho, T. is at the Japan Biological Information Research Center, Koto-ku, Tokyo, 135-0064 Japan		
	Yamagishi, A. is at Tokyo University of Pharmacy and Life Science, Hachioji, Tokyo, 192-0392 Japan		
	Nishio, Y., Usuda, Y. and Sugimoto, S. are at the Ajinomoto Co., Inc., Kawasaki, Kanagawa, 210-8681 Japan		
	The other authors are at the National Institute of Technology and Evaluation, Shibuya-ku, Tokyo, 151-0066 Japan.		
FEATURES	Location/Qualifiers		
source	1..300750		
	/organism="Corynebacterium efficiens YS-314"		
	/mol_type="genomic DNA"		
	/strain="YS-314"		
	/db_xref="taxon:196164"		
	252..620		
	/note="CE0885"		
	/codon_start=1		
	/transl_table=11		
	/product="hypothetical protein"		
	/protein_id="BAC17695.1"		
	/db_xref="GI:23492723"		
	/translation="MDQAPSGCLVDCVHGIFVDTVPMLLPVGTSGNRKIVLDLDFN		
	HYFLNKPDRILTLPLMSGCGCLGVWIGLTFPHWTHVYVCVGFDPGSGRFLAAC		
	TLHASRTMKPCLQGLVANG"		
	529..1936		
gene			


```

/ gene="rRNA_16S"
529..11936
/ gene="rRNA_16S"
/ product="16S ribosomal RNA"
/ protein_id="BAC17700.1"
/ note="CE-RNA01"
complement(2042..2215)
/ note="CE0886"
/ codon_start=1
/ transl_table=11
/ product="hypothetical protein"
/ protein_id="BAC17696.1"
/ db_xref="GI:23492724"
/ translation="MSSQTNSMPPTVHHHLGDVFADPWWMIHPIITCWQHPGNSNT
NNPHEHSPWWVY"
2399..5481
/ gene="rRNA_23S"
2399..5481
/ gene="rRNA_23S"
/ product="23S ribosomal RNA"
/ note="CE-RNA02"
5606..5725
/ gene="rRNA_5S"
5606..5725
/ gene="rRNA_5S"
/ product="5S ribosomal RNA"
/ note="CE-RNA03"
5925..7088
/ note="CE0889, similar to AX063829-1|CAC25156.1| percent
identity: 79 in 387 aa"
/ codon_start=1
/ transl_table=11
/ product="putative aminotransferase"
/ protein_id="BAC17699.1"
/ db_xref="GI:23492727"
/ translation="MRNKYVVERLRPFGETIFATMTORANEAGAINLQGGFPDGDGA
PMLEIASQQLGQNNQYSGGRGDGLRTAVASDRHKVPLKLDYDPTDLTVTGATEAI
SASVLGVLEVPYDEIVYDAAIAALAGATRAVPPLKEEDNTWVLDTDVHAHAV
TKSLMIIVNSPNCSTGVSFKSLGALAGIARAYDLLVLSDEYVHLLTDFDKVSHTA
AKTSGMDRTITVSSAKTENVTKGTGALAPAPILEAVIRAQKPFMSYVGATPQPA
VAHAHAKEPKWQMRGLONKRDILTALTOAGLKVDHSGTTFVVADIGERDGEAF
CFDLIDRVGVAIIPQAFVDPKPEQWSSKVPFACKQEDTLREAAQRLRAAGSL"
complement(7180..7821)
/ note="CE0890"
/ codon_start=1
/ transl_table=11
/ product="hypothetical protein"
/ protein_id="BAC17700.1"
/ db_xref="GI:23492728"
/ translation="MTHFSFDVDEAYARKKNEFLRDKRLQISAFFGLVLLAIGGVL
YVLADGAVGMVLIIVMGIMAAISFVMI PVLPROVGSASLNTYTPLVPAVVAEINPR
DVLLALVNTNVDETLPPRWGLAVRTINRLGVHDKLGERVPSVAVSGRTIRLDHWH
DEISPMPTTGWTTDADVVRRAEKTIPHDLWNKLEKNRRLDEVKNTRMNLLVL"
complement(7822..9495)
/ note="CE0891, similar to AL583924-93|CAC31112.1| percent
identity: 75 in 547 aa"
/ codon_start=1
/ transl_table=11
/ product="putative DNA helicase"
/ protein_id="BAC17701.1"
/ db_xref="GI:23492729"
/ translation="MAFGCDGLIIVQSDKTVLLETCHDLAGEARAALAPFAELERAPRH
VHTYRITPLALNARAAGHDAEQVLYSRFPVQPLLVDVAETMSRYGRVRMHK
HPAGLILESSEPAIIVLEISPRKIKMPLGOOLDPETIAVHPSERGLKQELLKIGWP
AEDLAGYVDEAHPIALSTEVDSLRDYQYAADSWEGSGVVVLPCCAGKTMVGA
ASMAQAQTLLITVNTVAGKQKDELLRTLTLEDEIGEYSERKKEIPVITATYQV
VTRTKGEYKALELFDPRDGLIITIDEVHLLPAPFRMTSLQSRRLGLTATLVRED
RGDGVFLSLGPKRDIAPWQKLEAQGIATACDCEVETRTDAAERMYATAETADRYR
LAATAHTKTRVRILOHQOPTLIIIGAYLDQLEELGAEFDAPIVDGKTSNKKRGEL
FDRFRSGLNVLVVSKVANFSIDLPAAVAIQVSGTFGSQEEAQRGLRLRPKADGG
EAHPYSVVSRTDLTDEVAHRQRFLAQGVAYRIIDADDLLFPITEKET"
complement(9548..11851)
/ note="CE0892, similar to AE006976-11|AAK45126.1| percent
identity: 36 in 741 aa"

```

```

/ codon_start=1
/ transl_table=11
/ product="conserved hypothetical protein"
/ protein_id="BAC17702.1"
/ db_xref="GI:23492730"
/ translation="MPCFPFSQIRQTKVTSTAPKLGIMKDSVPVTLTGWLEQLDDDD
QLSGILNRDPTALPLPPGLSAAARLQLRASIQRAVHKLSALELVALEAANAAGAE
NPITAEVVERLHRALHSAAPTAEQIGALNRLRLQALVFGGERTIILPETWATLP
TNQQLPEEGRSLTFGEASIEALPARHKILTVLTNGLGLTRDAALSADPARP
IPOLIKAGLLSRVDEQTVKLPGVVRVLEGEQDIAWVVPVPSAVGADAGIAAGL
EVARLQMLDALSTAPASTLKEGALGVNRWTRLTRELQVDEQOQVAVSLGVSSGLL
RRGVDPDLPADDGGDYIAPVVADEMLQATLAQLAHLRGMWTTQTYAALVGEADE
KRNPIHLSPASRRDALPETRAMILQSLTRVGEEDLEADLFFHRLAASRIPRETIRH
MLBEARWIGAVGGVTSPTARLTQAPAGVPEIADIADVAPKPDYFTVQADYTV
MVPCLQPEMQKIIGQLADESPGLASVYRISEASLRHAMDGLTATEIEOFLTOHSA
TGLPQSVAYILQDIARHGTLRGSPALCLSRDDPALLHSAVEAAGLGLRIAPTVA
VSNASLIEVITALRKAGMQPVAEDALGASLDLAPRPAPVPAEQPRAGPLDESIRA
AVAAIRREEVARTGTSDQTLAVLQAAVRGQRTVTLGFDVKQGVAVHRTVKPLTVNA
GQVDAIDESTGAVHRFMLEHRITEVID"
11836..12036
/ note="CE0893, similar to AL583924-91|CAC31110.1| percent
identity: 50 in 62 aa"
/ codon_start=1
/ transl_table=11
/ product="conserved hypothetical protein"
/ protein_id="BAC17703.1"
/ db_xref="GI:23492731"
/ translation="MGNMANVEKKYVDPAWPDHPDSEGHVVVTELIAPYACASSPWGD
DMEFPVPAETIGVHPYTRNR"
complement(12229..12822)
/ note="CE0894, similar to AX133781-1|CAC38957.1| percent
identity: 71 in 195 aa"
/ codon_start=1
/ transl_table=11
/ product="conserved hypothetical protein"
/ protein_id="BAC17704.1"
/ db_xref="GI:23492732"
/ translation="MLFMRHSKATTPFVKMAATTVAFGAAVAFAPAAASAAPDSW
DRLANCSGGNAINTCNGYHGGLQFSPSTWRYAGDFEFPAYQATREQOIIAUAERT
LAGQMGWAPACSSKGLNSAPTRNLNPAPAPAPAPAPAEYRATVDTNTNPVVGSS
DLNTVQGIIDAVSGTLAQYGVQVPAEIQAHYNAFGR"
13342..13725
/ note="CE0895, similar to AE006977-4|AAK45135.1| percent
identity: 52 in 127 aa"
/ codon_start=1
/ transl_table=11
/ product="putative cold shock protein"
/ protein_id="BAC17705.1"
/ db_xref="GI:23492733"
/ translation="MPVGTWKYDPERGFGFVSNPGGEDCFVGKQVLPKGVTELHQGQ
RIEFDPAAGRKGPQALRVKLETPTRRRPQHTYKPEELNGIISDMVTMLEGTVQALLK
GHYPRKVGQVAKILRAVAKELEA"
complement(13881..14408)
/ note="CE0896"
/ codon_start=1
/ transl_table=11
/ product="hypothetical protein"
/ protein_id="BAC17706.1"
/ db_xref="GI:23492734"
/ translation="MASRKTQRKNFIQIATLLVAVVVVVVIVVVFQNNNRPEPE
TVSITSSPAGIEVFPYSICEPGECAENEVPTLEVGADEELRTIPEI1HHDWYL
LTYVDDPAANDPEFYTTEVTVPGSDVPVTEGGERPRLVVVEVTSVMIGTDNDG
ETPVTTTSLATVEQ"
complement(14424..15254)
/ note="CE0897, similar to AE001874-1|AAF09704.1| percent
identity: 40 in 267 aa"
/ codon_start=1
/ transl_table=11
/ product="putative glutamine cyclotransferase"
complement(15255..15851)
/ note="CE0898, similar to AE001874-1|AAF09704.1| percent
identity: 40 in 267 aa"

```

Query Match 65.7%; Score 1036.4; DB 1; Length 300750;
 Best Local Similarity 81.1; Pred. No. 9.8e-186;
 Matches 1235; Conservative 0; Mismatches 246; Indels 41; Gaps 1;

```
QY 96 CATTGTGATCAGGTGAATAATCTAGTTAGCTCCCAAGTTGGCATAGGAGGCCACAGTGGC 155
Db 170262 CTTTTTGTGCGGTGAACATATCTGGTTAGTTTCCCAAGTTGGCATAGGAGGCCACAGTGGC 170321

QY 156 TGAATCATGCAGCTATTTCGCTCCGCGAAATTTCTCGACTCCCGCGGTAAACCCAAACCGTCGA 215
Db 170322 TGAATCATGCAGCTATTTCGCTCCGCGAGATCATGGATTCCCGCGGTAAACCCGACCGTCGA 170381

QY 216 GGCAGAGGTTTTCTCGATGACGGTTCCACGGTGTCCGAGGTGTCGCAAGGTGTTCCATCCGCGCATC 275
Db 170382 GGCAGAGGTTTTCTCGATGACGGTTCCACGGGTGTCGCGCGGTTCCTCCGAGCGTC 170441

QY 276 CACCGCGTCCAGAGGCTCATGAGCTCGGTGAGCGGTGGCGATCGCTACCTGGGCAAGG 335
Db 170442 CACCGGTGTCCAGAGGCGCACGAGCTGCGGTGACGGCGGAGATGTTACCTGGGCAAGG 170501

QY 336 CGTTTTGAAGCGAGTTGAACACGTCAACGAAGAAATCGGCGACGAGTCGCTGGCCTAGA 395
Db 170502 TGTCTCAAGGCTGTTGAGAACGTCACAGGAGATCGGCGACGAGCTGGCGCGCTGGA 170561

QY 396 GGTGACGATCAGGCGCTCATCGACGAAGCAATGATCAAGTTGATGGCACCGCCAAACAA 455
Db 170562 GGCAGACGATCAGGCTCTCATCGATGGCGGCATGATCAAGCTGGACGGCACCGGAGAACAA 170621

QY 456 GTCCCGCTCGGTGCAACCGCAATCTTGGTGTGTTCCATGGCTGTTGCAAGGCTGCTGC 515
Db 170622 GTCCCGCTCGGCGCCAAACGCCATCTTGGGTGTCTCCATGGCTGTGGCCAAAGGCGCGC 170681

QY 516 TGATTCGCGAGGCTCCACCTGTTCCGCTACATCGGTGGACCAAAACGACACAGCTTCTTCC 575
Db 170682 GGATTCGCGAGGCTCGCCCTGTTCCGTTACATCGGTGGCGCGAAACGCCACGCTGCTGCC 170741

QY 576 AGTTCGAATGATGAACATCATCAACGGTGGCGCTCAACGCTGACGTCCTCGGTGTTGACGTTCA 635
Db 170742 CGTCCCGATGATGAACATCATCAACGGTGGCGCACACCGCGACTCCCGGTGGATGTCCA 170801

QY 636 GGAATTCATGATCGCTCCAATCGGTGCAGACGCTTCTCGAGGCTCTCCGCAACGCGC 695
Db 170802 GGAATTCATGATCGCCCGCATCGGTTCGACTCTCTTCGAGGCGCTCGCGGAGGCGC 170861

QY 696 GGAGGTCTACACGCACTGAAGTCCGTCATCAAGGAAAGGCGCTGTCACCGGACTTGG 755
Db 170862 GGAGGTCTACACGCACTGAAGAGGTTCATCAACGAGAGGCGCTGTCACCGGTCTGG 170921

QY 756 CGATGAGGCGGCTTCGCTCTTCCGTCGCTCCACCGGTGAGGCTTTCGACCTTATCGT 815
Db 170922 CGATGAGGCGGCTTCGCTCCGTCGCTCCGTCGAGTCCACCGCTGACGCTTATCGT 170981

QY 816 TGAGGCAATCGAGAAGGCTGGCTTCACCCGAGGCAAGGACATCGCTCTTCTTGGACGT 875
Db 170982 CGAGGCCATCAAGAGCGCGCTTCGAGCGGGGCAAGGACATCGCTCTGCGCTGATGT 171041

QY 876 TGCTTCTCTGAGTTCTTCAAGACGCGACCTACCACTTCGAAGGTGGCCAGCACTCCGC 935
Db 171042 TGCTTCTCTGAGTTCTTCAAGAGCGGCAATACCACTTCGAGGTTGGCGAGCACCGC 171101

QY 936 AGCTGAGATGGCAACGTTTACGCTGAGCTCGTTGACGCGTACCCATCGTCTCCATCGA 995
Db 171102 CGAGGAGATGGCAACGCTTACGCGCGAGCTTGTGACGAGTACCCGATCGTCTCCATCGA 171161

QY 996 GGACCCACTCAGGAAGATGACTGGGAGGGTTACACCAACCTCACCGCAACCACTCGGCGA 1055
Db 171162 GGATCCGCTCAGGAGGACGACTGGGATGGTACGTCGACTGACCGCCGAGTCGGTGA 171221

QY 1056 CAAGGTTTCAGATCGTTGGCGACGACTTCTTCGTGTCACCAACCTTGAGCGCTGGAAGAGG 1115
Db 171222 CAAGGTTTCAGATCGTTGGCGATGACTTCTTCGTGTCACCAACCTTGAGCGCTGGAAGAGG 171281

QY 1116 CATCGCTTAAGAGGCTGCCAACTCCATCTGTTTAAAGTTGAACAGATCGGTACCTCAC 1175
Db 171282 CATCGCCAAAGAGGCGCCAACTCCATCTGTTTCAAGGTCAACAGATCGGTACCTCAC 171341
```

```
QY 1176 CGAGACCTTCGACGCTGTGCACATGGCTCACCGCGCAGGCTACACCTCCATGATGTCCCA 1235
Db 171342 CGAGACCTTCGACGCGCTGTGCACATGGCACACCGCGCAGGTTACACCTCCATGATGTCCCA 171401

QY 1236 CCGTTCGCGGTGAGACCGGAGGACACCACTTGTGTGACCTCGCAAGTTTCGACTCAACTGTGG 1295
Db 171402 CCGTTCGCGGTGAGACCGGAGGACACCACTTGTGTGACCTCGCGATCTCGGCTGTAAGTGGG 171461

QY 1296 CCAGATCAAGACTGGTGTCTCCAGCAGCTTCCGACCGGTTCGCAAGGTGTCGCAAGTTCCT 1355
Db 171462 CCAGATCAAGACCGGTGTCACCGCACGTTCCGACCGGTGGCCAAAGTACAAACAGTGTCT 171521

QY 1356 CCGATTCGACAGCTGCTTGGCGACCGCGGCTCTAGCAGGTGCGAGCGCATTCGCCAG 1415
Db 171522 GCGCATCGAGCAGCTGCTGGGTGACCGACGCGTCTAGCGCGCGGCTTCCCGGTTCCACG 171581

QY 1416 CTTTTCAGGGGTAAAA-----TAAAA 1434
Db 171582 TTTCCAGGCTAGACGCTTTTCAGAGNAACCTTGACCCCGCATCCCGTCACTTCTCAAA 171641

QY 1435 GCCTTTTTCGACCGCGGTAACTCAAGGTTCGCGGCGCTGTTGCTTACTACTGTTC 1494
Db 171642 AGTGGCGGATGTTGGCGTATCGGGCATGTTTCGGGCGTCGCAAGCAGTGGCGTGAC 171701

QY 1495 TGGTGTGACTATGATCGAGGATTATGCAAGAGCAGAGAAAATCTATAAAGGCTTGTTC 1554
Db 171702 GGGTGGGACTATGATGTTGATTATGGCGAAGCAGAGAAAACCCACAGGCGCATTTGTT 171761

QY 1555 CTGCTCAAGCAGGGAACGTGC 1576
Db 171762 CGGCTCTCAGCAGAGATCGTGC 171783

RESULT 13
BX248356 347625 bp DNA linear BCT 06-NOV-2003
LOCUS Corynebacterium diphtheriae gravis NCTC13129, complete genome;
segment 3/8.
ACCESSION BX248356 BX248353
VERSION BX248356.1 GI:38199583
KEYWORDS complete genome.
SOURCE Corynebacterium diphtheriae
ORGANISM Corynebacterium diphtheriae
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Corynebacteriaceae; Corynebacterium.
REFERENCE 1 (bases 1 to 347625)
AUTHORS Cerdeno-Tarraga, A.M., Efstathiou, A., Dover, L.G., Holden, M.T.G.,
Pallen, M., Bentley, S.D., Besra, G.S., Churcher, C., James, K.D., De
Zoyza, A., Chillingworth, T., Cronin, A., Dowd, L., Feltwell, T.,
Hamlin, N., Holroyd, S., Jagels, K., Moule, S., Quail, M.A.,
Rabinowitz, E., Rutherford, K., Thomson, N.R., Unwin, L.,
Whitehead, S. and Barrell, B.G. Parkhill, J.
The complete genome sequence and analysis of Corynebacterium
diphtheriae NCTC13129
Nucleic Acids Res. 31 (22), 6516-6523 (2003)
JOURNAL 14602910
PUBMED 2 (bases 1 to 347625)
REFERENCE Cerdeno-Tarraga, A.M.
AUTHORS Direct Submission
TITLE Submitted (03-OCT-2003) Cerdeno-Tarraga A.M., submitted on behalf
JOURNAL of the Pathogen Sequencing Unit, Sanger Institute, Wellcome Trust
Genome Campus, Hinxton, Cambridge CB10 1SA E-mail:
amct@sanger.ac.uk
FEATURES Location/Qualifiers
source 1. .347625
/organism="Corynebacterium diphtheriae"
/mol_type="genomic DNA"
/strain="NCTC13129"
/db_xref="taxon:1717"
/notice="biotype gravis"
/complement(81. .206)
/locus_tag="DIP0718"
complement(81. .206)
gene
CDS
```

misc_feature	gene	CDS	/locus_tag="DIP0718"	IMSRASVTL"
			/note="Doubtful CDS. No significant database matches"	
			/codon_start=1	
			/transl_table=11	
			/product="Putative membrane protein"	
			/protein_id="CAE49240.1"	
			/db_xref="GI:38199584"	
			/translation="MKYRSNIIIVCSLIPPQLPKARHIPAKIPSSPPVRIELP"	
			complement(141..206)	
			/locus_tag="DIP0718"	
gene	CDS		/note="Signal peptide predicted for DIP0718 by SignalP 2.0 HMM (Signal peptide probability 0.699) with cleavage site probability 0.667 between residues 22 and 23; signal-peptide site"	2250..5432 /locus_tag="DIP0722" 2250..5432 /locus_tag="DIP0722" /note="Similar to Streptomyces coelicolor putative ATP-dependent DNA helicase 2SCB6.07 TR:Q9FCK5 (EMBL:AL390968) (1159 aa) fasta scores: E(): 6.2e-29, 32.21% id in 1161 aa" /codon_start=1 /transl_table=11 /product="Putative ATP-dependent DNA helicase" /protein_id="CAE49244.1" /db_xref="GI:38199588" /translation="MDAFSTPQDSPTTSAQVIGEDLTAGIPPLQQSFVYAQRPRVEIV YNDHSEPIREWDMQFOEORGWRTVTGAAGVSTLTMDTVABRIRGVSPESIIVVA ASRESARLRAGIAHRIADGSYTPASLVSRVHSLAFALLRSLDEQLRLTGAEEVA VIRELQGHVDNFSLAAMPARQAQLGLIGFAREVDFLRSGERGLPEPDEELGA RYGRPMWASAGKFMREYQVMNIGGTHSNASSELVSKLCFDIPDMGWRITIIIDDAQH LDPQSAQLQALMRYTDFVTIADGPQQSFVFRGASPEFLHAPVDHELTYLSSFRQP TVEAKVLGSGDQSEFVADLLRRSHLLGIRWMDMAVIRSTADIPSLRRALLSAGVP VQEDPSDIIILSEORIVSALILAVRAIVQELSAQELELALGPICGADTVTLRLFRGL RKAEHRAGGNRAIEIMRELIDPKSESEQTQREQVEAVLTDRELAVVDKIRAVLVRG AQPSVEEILWEIWDASALSSHLOTVSLRGVGAQADRLDVAALFDDAAGDWVERR PTASITSFVRHIAEQELPTGVRDRRLTADAVRIVTAGHSLGQQMHTVI VAGVQEGTW PSIQETGTLFGQBDLIDLIDGIEPTNTIISRSAEKLEKRLFHVARTGTHRVVITA VESPEQDTAAEPSRFLQIGIAESKVQGPPTDPETASVQSEIDTFVHLDDQVYRLVSVPS IVAELERLANLPPSPORRQAARLARLARLALHQQVPGARPEQMGYGGPSETTSLDISK VPSLEIENALLCPLRAKLERLVEENTPIHMLKGTLAHAFAEAVGVGDVDPNEASQLVT QAFEALLDPAWSPHHMSQWMLQRLAHIDVSHAQRELVGVGVVNVVVPAGVEL RGRILRLERNDAGEFHIVDFKGTQKQAVTKDEANENKQLAYQLALHRKMLQRMGEPA INTTAEQGLTVDAQVILVYPATDKTIVTTREQAPKDAEELEKFSATLIPALLES LGPO LVARINTRCDCKIKTMCPAQPEGKMWPE"
			/locus_tag="DIP0720"	
			487..1365	
			/locus_tag="DIP0720"	
			/note="Similar to Mycobacterium tuberculosis hypothetical 31.0 kDa protein Rv3207c or MTCY07D11.19 TR:O05859 (EMBL:D95120) (285 aa) fasta scores: E(): 1.7e-44, 46.69% id in 287 aa"	
			/codon_start=1	
			/transl_table=11	
			/product="Putative membrane protein"	
			/protein_id="CAE49242.1"	
			/db_xref="GI:38199586"	
misc_feature	gene	CDS	/translation="MTDPAESFFVRFARDYGMRAYAIPVLAVITVWVLIDVFRTPAE TTTTATVGGAAPTATATSAQKGPDPARQNRPDITAITELPSGPEFTQKGEYRTVGNAG HAGKHDKHVFYTVIIVENGINTAAVGGDDAFAAMVDATLTNPKSWTHDKRGFEHV DAGAVKDPDLRIQLSSVDVTHGLCGNNIAMETSCFYGIGNRVINERWYRGAKPFG DLGAYRQYLINHEVGHIGIFANHEPCNGELAPIMMQQTLSLSNSELPAIDANETYN DDGAVCSANPWPYPFA"	2250..2558 /locus_tag="DIP0722" /note="ScanRegExp hit to PS00430, TonB-dependent receptor proteins signature 1." 3045..3200 /locus_tag="DIP0722" /note="HMPfam hit to PF00580, UvrD/REP helicase" 3285..3431 /locus_tag="DIP0722" /note="HMPfam hit to PF00580, UvrD/REP helicase" 4167..4199 /locus_tag="DIP0722" /note="ScanRegExp hit to PS00136, Serine proteases, subtilase family, aspartic acid active site." 5436..8656 /locus_tag="DIP0723" 5426..8656 /locus_tag="DIP0723" /note="Similar to Mycobacterium tuberculosis CDC1551 helicase, UvrD/Rep family MT3295 TR:AAK47638 (EMBL:AE007142) (1101 aa) fasta scores: E(): 2.9e-25, 43.03% id in 1120 aa" /codon_start=1 /transl_table=11 /product="Putative helicase" /protein_id="CAE49245.1" /db_xref="GI:38199589" /translation="MLSPQELSCALGKQFPPTQQADVISPLAFTLVVAGAGAKTE TMAARVWLVASGLVDPDRVLGLTFTTKAAQQLSKRIRDRLEQLAGIDNLRDLPTGA LATKLEAIPTVSTYDSYAGRLI SEYGLLIPVPSRLISQTELFQIASHIVSAHTGA LNTNSPNTVTSTLISLVSEMDNHMSPDTIEESAFLAMIEDVEATSKRAPSKEYI KWRDTQVIRNELLPLVQQLKTHLADNHLMTFGQMSLAARLAAENPQVGASQRRRYQI IMLDXYQDTGHAQORVLLKSLFAGTAVTAGVDPMQSIYGRGATAANLERLFTDFGSGI FPMDBERSVYADHAAQABQHDGHPFTAAVLVRKKHSAATALEQQGVPEVIEIVG LAGLIGIPEVADILVAJLTLVVRPYDTQAMRILAGPSVGLGMADLMALSDRAYNSGR DRRTATTELSDRPLERLKOIADTTPSDQDSIVGLAEVADLERDLSRSGYSAKSGS ERUTLAARLURYLRNLSNLSPLDFLBIERVFGIRTEVLOREDPRSDGATGAHLDR FAEVQDFSRIPGANISLLDLYSLAESENLEPEGVQVTAARVQLTTLVHKAGLEW QHVAVLHADANTYVAKASTLWTNASAVPSALRGDAKGDDELVGAPVFEIDTPTDAEAL
			/locus_tag="DIP0720"	
			/note="1 probable transmembrane helix predicted for DIP0720 by TMHMM2.0"	
			1165..1194	
			/locus_tag="DIP0720"	
			/note="ScanRegExp hit to PS00142, Neutral zinc metalloproteinases, zinc-binding region signature."	
			1389..2246	
			/locus_tag="DIP0721"	
			1389..2246	
			/locus_tag="DIP0721"	
misc_feature	gene	CDS	/note="Similar to Mycobacterium leprae hypothetical protein ML0818 TR:Q9CCG7 (EMBL:AL583919) (297 aa) fasta scores: E(): 1.2e-34, 39.47% id in 266 aa"	2250..5432 /locus_tag="DIP0722" 2250..5432 /locus_tag="DIP0722" /note="Similar to Streptomyces coelicolor putative ATP-dependent DNA helicase 2SCB6.07 TR:Q9FCK5 (EMBL:AL390968) (1159 aa) fasta scores: E(): 6.2e-29, 32.21% id in 1161 aa" /codon_start=1 /transl_table=11 /product="Putative ATP-dependent DNA helicase" /protein_id="CAE49244.1" /db_xref="GI:38199588" /translation="MDAFSTPQDSPTTSAQVIGEDLTAGIPPLQQSFVYAQRPRVEIV YNDHSEPIREWDMQFOEORGWRTVTGAAGVSTLTMDTVABRIRGVSPESIIVVA ASRESARLRAGIAHRIADGSYTPASLVSRVHSLAFALLRSLDEQLRLTGAEEVA VIRELQGHVDNFSLAAMPARQAQLGLIGFAREVDFLRSGERGLPEPDEELGA RYGRPMWASAGKFMREYQVMNIGGTHSNASSELVSKLCFDIPDMGWRITIIIDDAQH LDPQSAQLQALMRYTDFVTIADGPQQSFVFRGASPEFLHAPVDHELTYLSSFRQP TVEAKVLGSGDQSEFVADLLRRSHLLGIRWMDMAVIRSTADIPSLRRALLSAGVP VQEDPSDIIILSEORIVSALILAVRAIVQELSAQELELALGPICGADTVTLRLFRGL RKAEHRAGGNRAIEIMRELIDPKSESEQTQREQVEAVLTDRELAVVDKIRAVLVRG AQPSVEEILWEIWDASALSSHLOTVSLRGVGAQADRLDVAALFDDAAGDWVERR PTASITSFVRHIAEQELPTGVRDRRLTADAVRIVTAGHSLGQQMHTVI VAGVQEGTW PSIQETGTLFGQBDLIDLIDGIEPTNTIISRSAEKLEKRLFHVARTGTHRVVITA VESPEQDTAAEPSRFLQIGIAESKVQGPPTDPETASVQSEIDTFVHLDDQVYRLVSVPS IVAELERLANLPPSPORRQAARLARLARLALHQQVPGARPEQMGYGGPSETTSLDISK VPSLEIENALLCPLRAKLERLVEENTPIHMLKGTLAHAFAEAVGVGDVDPNEASQLVT QAFEALLDPAWSPHHMSQWMLQRLAHIDVSHAQRELVGVGVVNVVVPAGVEL RGRILRLERNDAGEFHIVDFKGTQKQAVTKDEANENKQLAYQLALHRKMLQRMGEPA INTTAEQGLTVDAQVILVYPATDKTIVTTREQAPKDAEELEKFSATLIPALLES LGPO LVARINTRCDCKIKTMCPAQPEGKMWPE"
			/locus_tag="DIP0720"	
			487..1365	
			/locus_tag="DIP0720"	
			/note="Similar to Mycobacterium tuberculosis hypothetical 31.0 kDa protein Rv3207c or MTCY07D11.19 TR:O05859 (EMBL:D95120) (285 aa) fasta scores: E(): 1.7e-44, 46.69% id in 287 aa"	
			/codon_start=1	
			/transl_table=11	
			/product="Putative membrane protein"	
			/protein_id="CAE49242.1"	
			/db_xref="GI:38199586"	
misc_feature	gene	CDS	/translation="MTDPAESFFVRFARDYGMRAYAIPVLAVITVWVLIDVFRTPAE TTTTATVGGAAPTATATSAQKGPDPARQNRPDITAITELPSGPEFTQKGEYRTVGNAG HAGKHDKHVFYTVIIVENGINTAAVGGDDAFAAMVDATLTNPKSWTHDKRGFEHV DAGAVKDPDLRIQLSSVDVTHGLCGNNIAMETSCFYGIGNRVINERWYRGAKPFG DLGAYRQYLINHEVGHIGIFANHEPCNGELAPIMMQQTLSLSNSELPAIDANETYN DDGAVCSANPWPYPFA"	2250..5432 /locus_tag="DIP0722" /note="ScanRegExp hit to PS00430, TonB-dependent receptor proteins signature 1." 3045..3200 /locus_tag="DIP0722" /note="HMPfam hit to PF00580, UvrD/REP helicase" 3285..3431 /locus_tag="DIP0722" /note="HMPfam hit to PF00580, UvrD/REP helicase" 4167..4199 /locus_tag="DIP0722" /note="ScanRegExp hit to PS00136, Serine proteases, subtilase family, aspartic acid active site." 5436..8656 /locus_tag="DIP0723" 5426..8656 /locus_tag="DIP0723" /note="Similar to Mycobacterium tuberculosis CDC1551 helicase, UvrD/Rep family MT3295 TR:AAK47638 (EMBL:AE007142) (1101 aa) fasta scores: E(): 2.9e-25, 43.03% id in 1120 aa" /codon_start=1 /transl_table=11 /product="Putative helicase" /protein_id="CAE49245.1" /db_xref="GI:38199589" /translation="MLSPQELSCALGKQFPPTQQADVISPLAFTLVVAGAGAKTE TMAARVWLVASGLVDPDRVLGLTFTTKAAQQLSKRIRDRLEQLAGIDNLRDLPTGA LATKLEAIPTVSTYDSYAGRLI SEYGLLIPVPSRLISQTELFQIASHIVSAHTGA LNTNSPNTVTSTLISLVSEMDNHMSPDTIEESAFLAMIEDVEATSKRAPSKEYI KWRDTQVIRNELLPLVQQLKTHLADNHLMTFGQMSLAARLAAENPQVGASQRRRYQI IMLDXYQDTGHAQORVLLKSLFAGTAVTAGVDPMQSIYGRGATAANLERLFTDFGSGI FPMDBERSVYADHAAQABQHDGHPFTAAVLVRKKHSAATALEQQGVPEVIEIVG LAGLIGIPEVADILVAJLTLVVRPYDTQAMRILAGPSVGLGMADLMALSDRAYNSGR DRRTATTELSDRPLERLKOIADTTPSDQDSIVGLAEVADLERDLSRSGYSAKSGS ERUTLAARLURYLRNLSNLSPLDFLBIERVFGIRTEVLOREDPRSDGATGAHLDR FAEVQDFSRIPGANISLLDLYSLAESENLEPEGVQVTAARVQLTTLVHKAGLEW QHVAVLHADANTYVAKASTLWTNASAVPSALRGDAKGDDELVGAPVFEIDTPTDAEAL
			/locus_tag="DIP0720"	
			/note="1 probable transmembrane helix predicted for DIP0720 by TMHMM2.0"	
			1165..1194	
			/locus_tag="DIP0720"	
			/note="ScanRegExp hit to PS00142, Neutral zinc metalloproteinases, zinc-binding region signature."	
			1389..2246	
			/locus_tag="DIP0721"	
			1389..2246	
			/locus_tag="DIP0721"	

[illegible]

FEATURES	Location/Qualifiers	gene	CDS	stem_loop	gene	CDS	repeat_unit	repeat_unit
	Location/Qualifiers 1. .276800 /organism="Streptomyces coelicolor A3 (2)" /mol_type="genomic DNA" /strain="A3 (2)" /db_xref="taxon:100226" 48. .98 /note="Score 59: 21/22 (95%) matches, 0 gaps" complement(146. .652) /gene="SCO3054" /note="synonym: SCBAC19G2.09c" complement(146. .652) /gene="SCO3054" /note="SCBAC19G2.09c, conserved hypothetical protein, len: 168aa: similar to many in Streptomyces coelicolor eg. TR:Q9F2X0 (EMBL:AL392148) SCD20.09 hypothetical protein (190 aa) fasta scores; opt: 250, Z-score: 285.6, 34.503% identity (38.816% ungapped) in 171 aa overlap." /codon_start=1 /transl_table=11 /product="conserved hypothetical protein" /protein_id="CAC44519.1" /db_xref="GI:14970940" /db_xref="UniProt/TREMBL:Q93J49" /translation="MSERSPAGGIGLIVETILVNTLDLTGADSLDTEGRAPCLTEDIYPAAREKSRATLLAGHGHPRHVRPTLGEILAAAPLVAVDPADGSAALAPAGEGSLTARUEAAVAQALVAGTWSRLKACEAADCHWAYDRSPAGRWCSMQVCGARAKM RYRAKES" 764. .1123 /gene="SCO3055" /note="synonym: SCBAC19G2.10" 764. .1123 /gene="SCO3055" /note="SCBAC19G2.10, conserved hypothetical protein, len: 119aa: similar to many in Streptomyces coelicolor e.g. TR:O95701 (EMBL:AF072709) 3SCF60.11c conserved hypothetical protein (131 aa) fasta scores; opt: 322, Z-score: 387.9, 42.017% identity in 119 aa overlap and downstream neighbouring CDS SCBAC19G2.11" /codon_start=1 /transl_table=11 /product="conserved hypothetical protein" /protein_id="CAC44520.1" /db_xref="GI:14970941" /db_xref="UniProt/TREMBL:Q93J48" /translation="MPVAELGPVVLDGCDPRALARFYADVILGTVBEGDWDVLRPLGGRAVAFQAPGFPVPPKMPAPDSSQOQFHLDLVDKDLDAEAKVLELGARPLDAEDRTGT FRVADPAGHPFCLICAC" 786. .1114 /gene="SCO3055" /note="Directly repeated at 11832. .12181" 1153. .1545 /gene="SCO3056" /note="synonym: SCBAC19G2.11" 1153. .1545 /gene="SCO3056" /note="SCBAC19G2.11, conserved hypothetical protein, len: aa: similar to many in Streptomyces coelicolor eg. TR:O9F3E9 (EMBL:AL450289) SC2H2.06 hypothetical protein (150 aa) fasta scores; opt: 340, Z-score: 415.4, 46.032% identity in 136 aa overlap and upstream neighbouring CDS SCBAC19G2.10" /codon_start=1 /transl_table=11 /product="conserved hypothetical protein" /protein_id="CAC44521.1" /db_xref="GI:14970942" /db_xref="UniProt/TREMBL:Q93J47" /translation="MAPAARFRSVVVDGPDPRILARFYAAVGGTTPDEADPDWVVLQVPGGPRLPADPLTTPWPRSRDRAQOQFHLDFDGGATWAEWMDAAHDRVIALGARPLD LEDREDKDFWYADPAGHPFCLCRIEHT" 1178. .1527 /gene="SCO3056"							
	/note="Directly repeated at 11440. .11768" complement(1558. .2688) /gene="SCO3057" /note="synonym: SCBAC19G2.12c" complement(1558. .2688) /gene="SCO3057" /note="SCBAC19G2.12c, possible dipeptidase, len: 376 aa: similar to many eukaryotic peptidases e.g. SW:P31430 (MDPI_RAT) rat renal dipeptidase (410 aa) fasta scores; opt: 440, Z-score: 474.0, 36.056% identity (38.906% ungapped) in 355 aa overlap. Also weakly similar to TR:O9KH70 (EMBL:AF288476) thermostable dipeptidase from Brevibacillus borstelensis (307 aa) fasta scores; opt: 273, Z-score: 298.4, 28.159% identity (30.116% ungapped) in 277 aa overlap. Contains Pfam match to entry PF01244 Renal dipeptase, Renal dipeptidase. Also similar to neighbouring upstream CDS SCBAC19G2.13c" /codon_start=1 /transl_table=11 /product="putative dipeptidase" /protein_id="CAC44522.1" /db_xref="GI:14970943" /db_xref="GOA:Q93J46" /db_xref="UniProt/TREMBL:Q93J46" /translation="MADLQDDLHTGAGGLDPLVLAEEAPFPFYPVSPDDDEP LARAHVAHPVADGNGPLPWALRHLPWYDLRGLBSAVDTVPRLREGHVGALLWSL HLPESLDGPRVAGVATLQOLDVKTVRVRAHPEGLRLAYDAGQAIADRMCGRIAVLPQGA GAALGRDLGILRSLHALGLRVLTLQGVSWASEAGLTRGEVVRNRLGVVADLSG ASATVRRTPAVSKAPALCTRSAARALRPHANLPDOLLVELGAAGLGNVPLTASQT GPVTRVDADLHVRTVAGVQSVGLSGTYDVGAAHPELGDPSCYPRLVAELLRRGMD EADVALLTGNVQVRFAAAFTAKAAQLRRREPSTATIADLDG" complement(1561. .2616) /gene="SCO3057" /note="Pfam match to entry PF01244 Renal dipeptase, Renal dipeptidase, score 121.70, E-value 1.4e-32" complement(1562. .2643) /note="Degenerately, directly repeated at 13490. .14610" complement(2808. .4010) /gene="SCO3058" /note="synonym: SCBAC19G2.13c" complement(2808. .4010) /gene="SCO3058" /note="SCBAC19G2.13c, possible dipeptidase, len: 400 aa: similar to many eukaryotic peptidases e.g. SW:P31429 (MDPI_RABIT) rabbit renal dipeptidase (410 aa) fasta scores; opt: 788, Z-score: 879.1, 42.432% identity (48.58% ungapped) in 403 aa overlap. Also weakly similar to TR:O9KH70 (EMBL:AF268476) thermostable dipeptidase from Brevibacillus borstelensis (307 aa) fasta scores; opt: 450, Z-score: 506.1, 31.034% identity (36.486% ungapped) in 348 aa overlap. Contains Pfam match to entry PF01244 Renal dipeptase, Renal dipeptidase. Also similar to neighbouring downstream CDS SCBAC19G2.12c" /codon_start=1 /transl_table=11 /product="putative dipeptidase" /protein_id="CAC44523.1" /db_xref="GI:14970944" /db_xref="GOA:Q93J45" /db_xref="UniProt/TREMBL:Q93J45" /translation="MTSLEKARELLRFFVVDGHNLDLPWALREQVRYDLDARDIADQ SAHLHTDLARLGRSGVGAOVSVVRSRDLPGVATLQIDCVRRLLIDRHPGELRAAL TAADMEAAREGRIASLIMGAGGHSIDNSLATLRLALYGLVRYMTLTHDNNWADSA TDEPGVGLSAFGREVVRNRMGLVLSHVAATMRDALDTSTAPVIFSHSSSRVAF EAMKRVAAFEERVPVRVATVSTVADHLDMREVGVDHLGIDGYDGTPTPDGLGDV SGYPNLITAEILLDRGWSQSLAKLTWKNVAVRVLDAAEDVSRGLRAARGPSNATIEQLDG TAAEQPEG" complement(2820. .4010) /gene="SCO3058" /note="Pfam match to entry PF01244 Renal dipeptase, Renal dipeptidase, score 280.40, E-value 2.3e-80" complement(2836. .3956)							
		misc_feature						
		repeat_region						
		gene						
		CDS						
		misc_feature						
		repeat_region						

FEATURES	Location/Qualifiers	gene	CDS	stem_loop	gene	CDS	repeat_unit	repeat_unit
	1. .276800 /organism="Streptomyces coelicolor A3 (2)" /mol_type="genomic DNA" /strain="A3 (2)" /db_xref="taxon:100226" 48. .98 /note="Score 59: 21/22 (95%) matches, 0 gaps" complement(146. .652) /gene="SCO3054" /note="synonym: SCBAC19G2.09c" complement(146. .652) /gene="SCO3054" /note="SCBAC19G2.09c, conserved hypothetical protein, len: 168aa: similar to many in Streptomyces coelicolor eg. TR:Q9F2X0 (EMBL:AL392148) SCD20.09 hypothetical protein (190 aa) fasta scores; opt: 250, Z-score: 285.6, 34.503% identity (38.816% ungapped) in 171 aa overlap." /codon_start=1 /transl_table=11 /product="conserved hypothetical protein" /protein_id="CAC44519.1" /db_xref="GI:14970940" /db_xref="UniProt/TREMBL:Q93J49" /translation="MSERSPAGGIGLIVETILVNTLDLTGADSLDTEGRAPCLTED DYPAAERKESRATLLAGHGHPRHVRPTLGEILAAAPLVAVDPADGSAALAPAGE GSLTARVAAVAQALVALVAGTWSRLKACEAADCHWAYDPSAGRWCMSQVCGARAKM RYRAKES" 764. .1123 /gene="SCO3055" /note="synonym: SCBAC19G2.10" 764. .1123 /gene="SCO3055" /note="SCBAC19G2.10, conserved hypothetical protein, len: 119aa: similar to many in Streptomyces coelicolor e.g. TR:O95701 (EMBL:AF072709) 3SCF60.11c conserved hypothetical protein (131 aa) fasta scores; opt: 322, Z-score: 387.9, 42.017% identity in 119 aa overlap and downstream neighbouring CDS SCBAC19G2.11" /codon_start=1 /transl_table=11 /product="conserved hypothetical protein" /protein_id="CAC44520.1" /db_xref="GI:14970941" /db_xref="UniProt/TREMBL:Q93J48" /translation="MPVAELGPVVLDGCDPRALARFYADVILGTVBEGDWDVLRPLG GRALAFQAPGFPVPPKPPAPDSSQOQFHLDLVDKDLDAEAKVLELGARPLDAEDRTGT FRVADPAGHPFCLICAC" 786. .1114 /gene="SCO3055" /note="Directly repeated at 11832. .12181" 1153. .1545 /gene="SCO3056" /note="synonym: SCBAC19G2.11" 1153. .1545 /gene="SCO3056" /note="SCBAC19G2.11, conserved hypothetical protein, len: aa: similar to many in Streptomyces coelicolor eg. TR:O9F3E9 (EMBL:AL450289) SC2H2.06 hypothetical protein (150 aa) fasta scores; opt: 340, Z-score: 415.4, 46.032% identity in 136 aa overlap and upstream neighbouring CDS SCBAC19G2.10" /codon_start=1 /transl_table=11 /product="conserved hypothetical protein" /protein_id="CAC44521.1" /db_xref="GI:14970942" /db_xref="UniProt/TREMBL:Q93J47" /translation="MAPAARFRSVVVDGPDPRILARFYAAVGGTGPDEADPDWVVLQV PGGPRLGFORAPDLTPPEWRPSDRNAQOQFHLDFDGGATWAEWMDAAHDRVLALGARPLD LEDREDKDFWYADPAGHPFCLCRIEHT" 1178. .1527 /gene="SCO3056"							
	/note="Directly repeated at 11440. .11768" complement(1558. .2688) /gene="SCO3057" /note="synonym: SCBAC19G2.12c" complement(1558. .2688) /gene="SCO3057" /note="SCBAC19G2.12c, possible dipeptidase, len: 376 aa: similar to many eukaryotic peptidases e.g. SW:P31430 (MDPI_RAT) rat renal dipeptidase (410 aa) fasta scores; opt: 440, Z-score: 474.0, 36.056% identity (38.906% ungapped) in 355 aa overlap. Also weakly similar to TR:O9KH70 (EMBL:AF288476) thermostable dipeptidase from Brevibacillus borstelensis (307 aa) fasta scores; opt: 273, Z-score: 298.4, 28.159% identity (30.116% ungapped) in 277 aa overlap. Contains Pfam match to entry PF01244 Renal dipeptase, Renal dipeptidase. Also similar to neighbouring upstream CDS SCBAC19G2.13c" /codon_start=1 /transl_table=11 /product="putative dipeptidase" /protein_id="CAC44522.1" /db_xref="GI:14970943" /db_xref="GOA:Q93J46" /db_xref="UniProt/TREMBL:Q93J46" /translation="MADLQDDLHTGAGGLDPLVLAEEAPFPFYPVSPDDDEP LARAHVAHPVADGNGPLPWALRHLPWDLRLGHSADVTDVLRREGHVGALLWSL HLPESLDGPRVAGVATLQOLDLVKTVRAHPEGLRLAYDAGQADARNCGRIAVLPQGA GAALGRDLGILRSLHALGLRVLTLQGVSWASEAGLTRGEEVVRNRLGVVADLSG ASATVRRTPAVSKAPALCTRSAARALRPHANLPDOLLVELGAAGLGNVPLTASQT GPVTRVDADLHVRTVAGPSVGLSGTGYDVGAAHPELGDPSCYPRLVAELLRRGMD EADVALLTGNVQVRLFAAAFTAKAAQLRRREPSTATIADLDG" complement(1561. .2616) /gene="SCO3057" /note="Pfam match to entry PF01244 Renal dipeptase, Renal dipeptidase, score 121.70, E-value 1.4e-32" complement(1562. .2643) /note="Degenerately, directly repeated at 13490. .14610" complement(2808. .4010) /gene="SCO3058" /note="synonym: SCBAC19G2.13c" complement(2808. .4010) /gene="SCO3058" /note="SCBAC19G2.13c, possible dipeptidase, len: 400 aa: similar to many eukaryotic peptidases e.g. SW:P31429 (MDPI_RABIT) rabbit renal dipeptidase (410 aa) fasta scores; opt: 788, Z-score: 879.1, 42.432% identity (48.58% ungapped) in 403 aa overlap. Also weakly similar to TR:O9KH70 (EMBL:AF268476) thermostable dipeptidase from Brevibacillus borstelensis (307 aa) fasta scores; opt: 450, Z-score: 506.1, 31.034% identity (36.486% ungapped) in 348 aa overlap. Contains Pfam match to entry PF01244 Renal dipeptase, Renal dipeptidase. Also similar to neighbouring downstream CDS SCBAC19G2.12c" /codon_start=1 /transl_table=11 /product="putative dipeptidase" /protein_id="CAC44523.1" /db_xref="GI:14970944" /db_xref="GOA:Q93J45" /db_xref="UniProt/TREMBL:Q93J45" /translation="MTSLEKARELLRFFVVDGHNLDLPWALREQVRYDLDARDIADQ SAHLHTDLARLGRSGVGAOVSVVRSRDLPGVATLQIDCVRRLLIDRHPGELRAAL TAADMEAAREGRIASLIMGAGGHSIDNSLATLRLALYGLVRYMTLTHDNNWADSA TDEPGVGLSAFGREVVRNRMGLVLSHVAATMRDALDTSAPVIFSHSSSRVAF EAMKRVAAFEERVPVRVATVSTVADHLDMREVGVDHLGIGDYDGTPTPDGLGDV SGYPNLITAEILLDRGWSQSLAKLTWKNVAVRVLDAAEDVSRGLRAARGPSNATIEQLDG TAAEQPEG" complement(2820. .4010) /gene="SCO3058" /note="Pfam match to entry PF01244 Renal dipeptase, Renal dipeptidase, score 280.40, E-value 2.3e-80" complement(2836. .3956)							

gene /note="Degenerately, directly repeated at 12216. .13297"
complement(4025. .4567)
/gene="SCO3059"
CDS complement(4025. .4567)
/gene="SCO3059"
/EC number="4.1.1.21"
/note="SCBAC19G2.14c, phosphoribosylaminimidazole
carboxylase catalytic subunit PurE, len: 180aa: strongly
similar to many eg. SW:044679 (PurE_CORAM)
phosphoribosylaminimidazole carboxylase catalytic subunit
PurE from Corynebacterium ammoniagenes (177 aa) fasta
scores; opt: 666, Z-score: 735.8, 65.541% identity
(65.541% ungapped) in 148 aa overlap and TR:080937
(EMBL:AC004684) putative phosphoribosylaminimidazole
carboxylase from Arabidopsis thaliana (645 aa) fasta
scores; opt: 644, Z-score: 703.2, 60.870% identity
(60.870% ungapped) in 161 aa overlap. Contains Pfam match
to entry PF00731 AIRC, AIR carboxylase."
/codon_start=1

Query Match 46.5%; Score 734; DB 1; Length 276800;
Best Local Similarity 72.6%; Pred. No. 1.3e-128;
Matches 963; Conservative 0; Mismatches 360; Indels 3; Gaps 1;
QY 106 CGTGAATAATCTAGTCCCAAGTTGGCATAGGAGGCCACAGTGGCTGAAATCATG 165
DB 43666 CGTGAAGCGTCACAGTGAAGGCTCACATTTGAAGAGAGATGCTGTGCGCTCATCGAC 43607
QY 166 CAGCTATTTCGCTCGCGAAATTTCTGACTCCCGCGGTAAACCAACCGTCGAGGCGAGAGTT 225
DB 43606 GTGCTGTAGCCGGGAAATCTTGGACTCCCGAGGCAACCCACGGTCGAGGTCGAGGTC 43547
QY 226 TTCCTGATGACGTTCCACGGTGTGCGAGTGTTCATCCCGCGCATCCACGGCGTC 285
DB 43546 GGCCTCGACGACGCGACGCGGTGTGCGCGCGTTCGCTCGCGCGCTCCACCGCGCC 43487
QY 286 CACGAGGCTCATGAGTTCGTCGAGTGGC---GATCGTACCTGGGCAAGGGGCTTTG 342
DB 43486 TTCGAGGCGATCGAGTTCGTCGAGCGGACCCGAGCGGTACCTCGGCAAGGGGCTCGAG 43427
QY 343 AAGCAGTTGAAACGTCACAGAAATCGCGACAGCTCGCTGCGCTAGAGGCTGAC 402
DB 43426 AAGCCGTGTGCGCGTCATCGACAGATCGGCCGCGAGTGTGTCGCTACGACGCCACC 43367
QY 403 GATCAGCGCTCATCGAGCAAGCAATGATCAAGCTTGATGSCACCGCCCAACAAGTCCGC 462
DB 43366 GAGCAGCGCTGATCGACCAAGCGCATGTCACCTGGACGCGCACCAAGGGCTCG 43307
QY 463 CTGGGTGCAACGCAATCTTGGTGTTCATGCTGTTCGAAAGGCTGCTGATTCC 522
DB 43306 CTGCGCGCAACGCCATCTCTCGGCTGTCTCTCGCGTCCGCGCGCTCCGAGGCC 43247
QY 523 GCAGGCTTCCCATGTTTCGCTACATCGTGGACCAACGACACAGTTCCTTCCAGTTCCA 582
DB 43246 AGCGACTGCGCTCTTCCGCTACTGTCGGCGCGCGAACGCGACCTGCTCGCGTGGCG 43187
QY 583 ATGATGAACATCATCAAGGTGGCGCTCACGCTGACTCCGCTGTGAGTTTCAGGAATTC 642
DB 43186 ATGATGAACATCTTGAACGCGCGCTCGCACGCCACTCCAAAGTGGACATCCAGGAGTTC 43127
QY 643 ATGATCGCTCAATCGGTGAGAGACCTTCTCTGAGGCTCTCCGCAACGCGCGGAGGTC 702
DB 43126 ATGATCGCGCGATCGGCGGAGTCTCTTTCGAGGCGCTGCGCTGGGCGCGAGGTC 43067
QY 703 TACCACGACATGAAGTCGTCATCAAGAAAGGCGCTGTCCACCGGACTTGGCGATGAG 762
DB 43066 TACCACACCTCAAGAAAGTCTCTGAAGAAACAGGCGCTGGCACCGGCTCGGCGACGAG 43007
QY 763 GCGGCTTCGCTCTTCCGTCGCTCCACCGTGGAGGCTTTGACCTTATCTTGAAGCA 822
DB 43006 GCGGCTTCGCGCGCAACCTGGGCTCCAAACCGGAGGCGCTCGACCTCATCTCGAGGCG 42947

QY 823 ATCGAGAAGGCTGGCTTACCCCGAGCAAGGACATCGCTTGTCTGTGAGACGTGTCTTC 882
DB 42946 ATCAAGGAGCGGCTACACCCCGGCGGAGCAGATCGCCCTCGGCTCGAGCTGCGCG 42887
QY 883 TCTGAGTTTCAAGGACGGCACTACCACTTCAAGGTTGGCGAGCACTCCGAGCTGAG 942
DB 42886 TCCGAGTTCTACAGGACGGCTCTACGCCCTTCGAGGGCAAGAACCGCTCCGCGCGAG 42827
QY 943 ATGGCAAAAGCTTACCGTGAAGTGTGAGCGGTACCAATCTCTCCATCGAGGACCCA 1002
DB 42826 ATACCGAGTACCTACCGGAGCTGGTTCGAGGCGTACCCGCTGTCTCTCAATCGAGGACCCG 42767
QY 1003 CTGCAAGAGATGACTGGGAGGCTTACACCAACTCACCCCAACCACTACCGCGAGCAAGGTT 1062
DB 42766 CTGTTGAGGACGACTGGGAGGCTGGAACACCATCACCGCAAGCTCGGCGACAGGTG 42707
QY 1063 CAGATCGTTGGGACGACTTCTTGTGTCACCAACCTGAGCGCTGAAGGAGGGGATCGCT 1122
DB 42706 CAGCTGTGCGGACGACCTTCTGTCACCAACCGGAGCGCTTGGCGCGGATCGAG 42647
QY 1123 AAGAGGCTGCCAACTCCATCTGTTAAGGTGAACAGATCGGTACCTCACCGAGACC 1182
DB 42646 GAGAACTCGGCCAAACGCCCTGCTCGTCAAGGTCAACAGATCGGTTTCGCTGACCGAGACT 42587
QY 1183 TTGACGCTGTGACATGGCTCACCGCGAGGCTACACCTTCCATGATGTCCACCGTTCC 1242
DB 42586 CTGACGCGCTCGAGCTGGGCCAGCGCAACGGCTTCAAGTGCATGATGTCCACCGCTCC 42527
QY 1243 GGTGAGACCGAGGACCAACATTTGCTGTCACCAACCTGAGCGCTGAAGGAGGGGATC 1302
DB 42526 GGGGAGACCGAGGACGCTGACATCGCGGACCTGCGCGCTGCCCACTCGCGGCGAGATC 42467
QY 1303 AAGACTGGTCTCCAGCAGCTCCGACCGGTGTCGAAAGTACAAACAGGTTTCTCCGCATC 1362
DB 42466 AAGACCGCGCGCGCGCGCGCTCCGAGCGCTGCCAAGTACAAACAGCTGTGTCGCGATC 42407
QY 1363 GAGCAGCTGTGTCGCGACGCGCGGCTGTACGAGGTGCGAGCGATTTCCACGCTTTCAG 1422
DB 42406 GAGGAGATCTCGACGAGCGCGGCTGTACGCGCGCGAGCGCTTCCCGCTTCAAG 42347
QY 1423 GGCTAA 1428
DB 42346 GGCTGA 42341
RESULT 15
AE016822_17/c
WPCOMMENT
Sequence split into 26 fragments LOCUS AE016822 Accession AE016822
Fragment Name Begin End
AE016822_00 1 110000
AE016822_01 100001 210000
AE016822_02 200001 310000
AE016822_03 300001 410000
AE016822_04 400001 510000
AE016822_05 500001 610000
AE016822_06 600001 710000
AE016822_07 700001 810000
AE016822_08 800001 910000
AE016822_09 900001 1010000
AE016822_10 1000001 1110000
AE016822_11 1100001 1210000
AE016822_12 1200001 1310000
AE016822_13 1300001 1410000
AE016822_14 1400001 1510000
AE016822_15 1500001 1610000
AE016822_16 1600001 1710000
AE016822_17 1700001 1810000
AE016822_18 1800001 1910000
AE016822_19 1900001 2010000
AE016822_20 200001 2110000
AE016822_21 2100001 2210000
AE016822_22 2200001 2310000

AE016822_23		2300001	2410000		
AE016822_24		2400001	2510000		
AE016822_25		2500001	2584158		
Continuation (18 of 26) of AE016822 from base 1700001 (AE016822 Leifsonia xyli subsp. xy)					
Query Match		44.5%	Score 701.6;	DB 1;	Length 110000;
Best Local Similarity		72.0%;	Pred. No. 1.8e-122;		
Matches 930;		Conservative 0;	Mismatches 359;	Indels 3;	Gaps 1;
Qy	140	AGGAGGCGACAGTGGCTGAAATCATGACGCTATTTCGTCGCGAAATTCCTCGACTCCCGCG	199		
Db	83109	AGGAGACTGTGTGGCTGCAATCGAAGCTGTAGGAGCTCGCGAGATCTCTCGACTCGCGCG	83050		
Qy	200	GTAACCCAAACGCTCGAGGACAGGTTTTCTTGATGACGGTTTCCACCGGTGTGCGAGGTG	259		
Db	83049	GAACACCGACCGTTCGAGTTCGAGTGTCTCTGAAAGATGGCAGCGTTCAGCGCGCGCGCG	82990		
Qy	260	TTCCATCCGGCGCATCCACCGGCGTCCACGAGGCTCATGAGCTCGGTGACGGTGGC---G	316		
Db	82989	TCCCGTCCGGCGGCTCCACCGGCGGTTTCGAGGCTTACGAGCTCCGCGACGCGACAGG	82930		
Qy	317	ATCGCTACTCTGGCAAGGGCGTTTTGAAGGCAGTTGMAAACGTCACAGAAATAATCGGCG	376		
Db	82929	GCCGCTACTCTGGCAAGGGTTCGAGAAGGCCGTGGATGCCGTGCTCGACGAGATCGGCC	82870		
Qy	377	ACGAGCTCGCTGGCTAGAGGCTGACGATCAGCGCTCATCGACGAAGCAATGATCAAGC	436		
Db	82869	CGGCCATCTGAGGGGCTTCGAGGCCAGCGACAGCGCATCTGTCGATGAGGGCGATGATCGAAC	82810		
Qy	437	TTGATGGCACCGCCACAAAGTCCCGCTCGGTGCAACGCAATCTTGGTGTTCATGG	496		
Db	82809	TCGACGGCACCGCACAAAGAACGCGCTGGGCGCTTAACGCCATGCTCGGTGTGAGCTCG	82750		
Qy	497	CTGTGTCAAAGGCTGCTGCTGATTCGCGAGGCTCCCACTGTTCCGCTACATCGGTGGAC	556		
Db	82749	CAGTCGCAAGGCGCGCGCGACTCCGCGACCTGCGCGCTGTTCCGCTACTCGGCGGCC	82690		
Qy	557	CAAAACGACACGTTCTTCCAGTTCCAAATGATGAACATCATCAAGGTGGCGCTCACGCTG	616		
Db	82689	CGAACGGCGATGTCTTCCCGCTCCGATGATGAACATCATCAACGGCGCGCGCACCGCG	82630		
Qy	617	ACTCCGGTGTGAGCTTCAGGAATTCATGATCGCTTCCAAATCGGTGCGAGACCTTCTCTG	676		
Db	82629	ACACGGGCTCGACATCCAGGAGTTCTATGATTCCTCCGATCGGCGCGGAGACCTTCTCCG	82570		
Qy	677	AGGCTCTCCGCAAACGGCGCGAGGTCTACACGCACTGAAGTCCGTATCAAGGAAAGG	736		
Db	82569	AGGGCTTCGCTGGGGCGGTGAGACCTTACCACTCGCTCAAGGCGCTGCTGAAGTCGAAGG	82510		
Qy	737	GCCTGTCCACCGGACTTGGCGATGAGGCGGCTTCGCTTCCGTCGGCTCCACCGCTG	796		
Db	82509	GCCTGAAACACCGGCTTGGGCGACGAGGGCGGCTTCGCCCGCGAGCTGGAGCAACACCGG	82450		
Qy	797	AGGCTCTTGAACCTTATCTGTTGAGGCAATCGAGAAGGTGGCTTCAACCCAGGCAAGGACA	856		
Db	82449	CTGGCTCGACCTGATCGCCGAGGCCATCGAGAAGCGCGCTTCAACGTGCGCTCGCAGA	82390		
Qy	857	TCGCTCTTGCTCTGGAGCTGCTTCTCTGAGTTCTTCAAGGACGGCACCTACCACTTCG	916		
Db	82389	TCGCGCTCGGCTCGACGTGGCTTCCACCGAGTTCTTCGAGAACGGGCGTCTACCGGCTCG	82330		
Qy	917	AAGGTGGCCAGCACTCCGCGAGCTGAGATGCAAAAGTTTACGCTGAGCTCGTTGACGCGT	976		
Db	82329	AGGGCGAGGACCGCACCGCGCGCGAGATGAGCGCTTACTACCCGATCTCGGAACT	82270		
Qy	977	ACCCAATCTGCTCCATCGAGGACCCACTGCGAGGAGATGACTGGGAGGGTTACCAACAC	1036		
Db	82269	ACCGCTGGTCTCCATCGAGGACCCGCTTGGCGGAGGACGACTGGAGGGCTGGGCGCAC	82210		
Qy	1037	TCACCGCAACCATCGGGGACAAAGTTTCAGATCGTTGGCGACGACTTCTTCGTCAACAC	1096		
Db	82209	TGAACGCCGAGATCTGGCTCGACGCTGCAACTCGCTCGGCGATGACCTGTTGTCGTCAACAC	82150		

Search completed: September 28, 2005, 20:45:20
Job time : 6718 secs

Qy	1097	CTGAGCGCCTGAAGGAGGCGATCGCTAAGAAGGCTGCCAATCTCATCTCTGTTAAGGTGA	1156
Db	82149	CGAAGCGCCTGGCGCAGGGCATCCAGCAGAAGGCCGCGAACAGCATCTCTGTCAAGGTGA	82090
Qy	1157	ACGAGATCGGTACCTCTACCGAGACCTTTCGACGCTGTGACATGGCTCACCGCGAGGCT	1216
Db	82089	ACCAGATCGGCACCTTTCGACCGAGACGCTGGACGCGGTCTCGCTCGCCCGAGCGCGTA	82030
Qy	1217	ACACTCCATGATGTCACCGTTCCGTTCCGTTGAGACCGAGGACACCATTTCTGACCTCG	1276
Db	82029	TGACCGCGCTCTCTCGACCGCTCCGCGGAGACCGAGGACACCATCCCGACCTCG	81970
Qy	1277	CAGTTGCACTCAACTGTGGCCAGATCAAGACTGTGTGCTCCAGACAGTTTCCACCGTGTCTCG	1336
Db	81969	CCGTGCGCACCGATTTCGCGCCAGATCAAGACCGCGCGCCCGGCTCCCGAGCGCTCG	81910
Qy	1337	CAAGTACACACCGCTTCTCGGATCGAGCAGCTGCTTGGGAGCGCGGCGCTCTACGCGAG	1396
Db	81909	CGAAGTACACACAGCTGCTCGGATCGAGGAGGAGCTGGGCGAGGCGGCTCTCTACGCGG	81850
Qy	1397	GTGCGACGCGATTCCCGACGCTTTCAGGCGCTAA	1428
Db	81849	CCGTTTCGGCGTTTCCCGCTTACCAGGCGCTGA	81818

THIS PAGE BLANK (user)

THIS PAGE BLANK (user)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 28, 2005, 16:00:03 ; Search time 793.5 Seconds
(without alignments)
11772.340 Million cell updates/sec

Title: US-10-728-947-1
Perfect score: 1578
Sequence: 1 ggctgggatatgggtagt.....ctcaagcagggaacgtgctt 1578

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq_16Dec04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002as:*
7: Geneseqn2002bs:*
8: Geneseqn2003as:*
9: Geneseqn2003bs:*
10: Geneseqn2003cs:*
11: Geneseqn2003ds:*
12: Geneseqn2004as:*
13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1578	100.0	1578	4	AAF61696
2	1578	100.0	349980	5	AAH68527
3	1403.4	88.9	1405	10	ADD13322
4	1396.4	88.5	1398	4	AAH71395
5	1278	81.0	1278	8	ACC45311
6	1275	80.8	1275	5	AAH66050
7	1275	80.8	1299	8	ACC45312
8	1273.2	80.7	1278	8	ACC45314
9	1270.2	80.5	1299	8	ACC45313
10	908.6	57.6	1275	8	ACA29854
11	679.8	43.1	1269	13	ADS56353
12	647	41.0	1287	8	ACA38079
13	642.4	40.7	30159	4	AAS59564
14	642.4	40.7	30159	4	ACF64493
15	632.6	40.1	45190	6	ABX09144
16	632.6	40.1	110000	4	AAI96882_11
17	632.6	40.1	110000	4	AAI96883_11
18	632.4	40.1	1317	8	ACA38642
19	626.2	39.7	1290	8	ACA40436
20	584.6	37.0	349980	6	ABQ81846

c	21	584.6	37.0	349980	6	ABQ81847
	22	583.4	37.0	1344	8	ACA39628
	23	582	36.9	1269	13	ADS58570
	24	575.4	36.5	1269	13	ADS59709
	25	574	36.4	1275	13	ADS59832
	26	574	36.4	1281	13	ADT42160
	27	569.8	36.1	1278	13	ADT43135
	28	569	36.1	1272	13	ADT45911
	29	567.4	36.0	1260	13	ADS62752
	30	567.4	36.0	1269	13	ADS63235
	31	561.8	35.6	1272	13	ADS55775
	32	557	35.3	1278	13	ADT46979
	33	552	35.0	1290	13	ADT44740
	34	542	34.3	1281	8	ACA23750
	35	533.6	33.8	1335	11	ABD05845
	36	533	33.8	1347	11	ABD05554
	37	531.8	33.7	1290	4	AAS54197
	38	531.8	33.7	1290	8	ACA42456
	39	530.8	33.6	1281	13	ADT46255
	40	528	33.5	1287	13	ADS56507
	41	522.8	33.1	1278	13	AD849945
	42	522.8	33.1	1281	8	ACA25228
	43	522.8	33.1	1281	13	ADS55460
	44	517.4	32.8	1690	12	ADO59752
	45	515.2	32.6	1404	8	ACA45728

ALIGNMENTS

RESULT 1

AAF61696
ID AAF61696 standard; DNA; 1578 BP.

XX AAF61696;

XX 12-JUL-2001 (first entry)

DE C. glutamicum enolase encoding DNA.

XX Enolase; fermentation; L-amino acid; L-lysine; coryneform; eno gene;
KW Enolase; animal feed supplement; ds.

XX Corynebacterium glutamicum.

PH Key Location/Qualifiers
FT CDS 151..1428

FT /*tag= a
FT /product= "Enolase"

XX EPI090998-A1.

XX PD 11-APR-2001.

XX 29-SEP-2000; 2000EP-00121158.

PR 05-OCT-1999; 99DE-01047791.

XX (DEGS) DEGUSSA-HUELS AG.

XX Moeckel B, Pfeifferle W, Hermann T, Puehler A, Kalinowski J;
XX Bathe B;

XX WPI; 2001-292928/31.

XX P-PSDB; AAB70882.

XX New enolase gene from coryneform bacteria, used to prepare transformants
XX with increased synthesis of amino acids, particularly lysine.

XX Claim 4; Page 12-14; 25pp; German.

XX This invention describes a novel isolated nucleic acid (I) from
XX coryneform bacteria which is used in a method for fermentative production

PR 16-DEC-1999; 99JP-00377484.
 PR 07-APR-2000; 2000JP-00159162.
 PR 03-AUG-2000; 2000JP-00280988.
 XX (KYOW) KYOWA HAKKO KOGYO KK.
 XX Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
 PI Tateishi N, Senoh A, Ikeda M, Ozaki A;
 XX WPI; 2001-376931/40.
 XX Novel polynucleotides derived from Coryneform bacteria, for identifying
 PT mutation point of a gene, measuring expression of a gene, analyzing
 PT expression profile or pattern of a gene and identifying homologous gene.
 XX Disclosure; SEQ ID NO 7062; 246pp + Sequence Listing; English.
 CC The present invention provides a number of nucleotide and protein
 CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These
 CC are useful for identifying the mutation point of a gene derived from a
 CC mutant of coryneform bacterium, measuring expression amount and analysing
 CC the expression profile or expression pattern of a gene derived from
 CC Coryneform bacterium, and identifying a homologue of a gene derived from
 CC coryneform bacterium. Coryneform bacteria are useful for producing amino
 CC acids, nucleic acids, vitamins, saccharides and organic acids,
 CC particularly L-lysine. The present sequence is a nucleic acid described
 CC in the exemplification of the invention. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from the European Patent Office
 XX
 SQ Sequence 349980 BP; 79725 A; 90426 C; 98918 G; 80911 T; 0 U; 0 Other;
 Query Match 100.0%; Score 1578; DB 5; Length 349980;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1578; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 GGCTGGGGATATGGTATGTTTCGCCACTAATTTCACTGATTCGCTCATCGAAACAAGA 60
 Db GGCTGGGGATATGGTATGTTTCGCCACTAATTTCACTGATTCGCTCATCGAAACAAGA 134858
 Qy 61 TTCTGTGCAAACTTGGGTGTAGAGTGATTTGAAGACATTTGATCAGTGAATATTTCTAG 120
 Db TTCTGTGCAAACTTGGGTGTAGAGTGATTTGAAGACATTTGATCAGTGAATATTTCTAG 134918
 Qy 121 TTAGTCCCAAGTTGGATAGAGGCCACAGTGGCTGAATCATGCAAGTATTCGTCGC 180
 Db TTAGTCCCAAGTTGGATAGAGGCCACAGTGGCTGAATCATGCAAGTATTCGTCGC 134978
 Qy 181 GAAATTTCTCGACTCCCGCGTAAACCCCAACCGTCGAGGCGAGAGGTTTCTCGATGACGGT 240
 Db GAAATTTCTCGACTCCCGCGTAAACCCCAACCGTCGAGGCGAGAGGTTTCTCGATGACGGT 135038
 Qy 241 TCCACAGGTGTGCGAGGTTCATTCGCGCGATCCACCGCGTCCAGAGGCTCATGAG 300
 Db TCCACAGGTGTGCGAGGTTCATTCGCGCGATCCACCGCGTCCAGAGGCTCATGAG 135098
 Qy 301 CTGCGTGAAGTGGGATCGCTACTCTGGCAAGGCGGTTTGAAGGCGATTTGAAGGCGTC 360
 Db CTGCGTGAAGTGGGATCGCTACTCTGGCAAGGCGGTTTGAAGGCGATTTGAAGGCGTC 135158
 Qy 361 AACGAAGAAATCGGCGACGAGTGGCTGGCTAGAGGCTGACGATCAGCGGCTCATCGAC 420
 Db AACGAAGAAATCGGCGACGAGTGGCTGGCTAGAGGCTGACGATCAGCGGCTCATCGAC 135218
 Qy 421 GAAGCAATGATCAAGCTTGTGATGGCAACCGCAAGTCCCGCTGGGTCGCAACCGCAATC 480
 Db GAAGCAATGATCAAGCTTGTGATGGCAACCGCAAGTCCCGCTGGGTCGCAACCGCAATC 135278
 Qy 481 CTGGTGTTCATGGCTGTTGCAAGGCTGCTGCTGATTCGGCAGGCTCCCACTGTTTC 540
 Db CTGGTGTTCATGGCTGTTGCAAGGCTGCTGCTGATTCGGCAGGCTCCCACTGTTTC 135338
 Qy 541 CGCTACATCGGTGGACCAACCGCACAGTTCCTTCCAGTTTCCAAATGATGAACATCATCAAC 600

Db CGCTACATCGGTGGACCAACCGCACAGTTCCTTCCAGTTCCAATGATGAACATCATCAAC 135338
 Qy 601 GGTGGCGCTCAGCTGACTCGGTGTTGAGTTTCAGAAATTCATGATCGCTCCAAATCGGT 660
 Db GGTGGCGCTCAGCTGACTCGGTGTTGAGTTTCAGAAATTCATGATCGCTCCAAATCGGT 135458
 Qy 661 GCAGAGACCTTCTCTGAGGCTCTCCGCAACGGCGCGAGGTCTACACGCACTGAAGTCC 720
 Db GCAGAGACCTTCTCTGAGGCTCTCCGCAACGGCGCGAGGTCTACACGCACTGAAGTCC 135518
 Qy 721 GTCATCAAGAAAAAGGCGCTGTCCACCGGACTTGGCGATGAGGCGGCTTCGCTCCTTCC 780
 Db GTCATCAAGAAAAAGGCGCTGTCCACCGGACTTGGCGATGAGGCGGCTTCGCTCCTTCC 135578
 Qy 781 GTGGGCTCCACCGCTGAGGCTCTTGACCTTATCGTTGAGCAATCGAGAGGCTGGCTTC 840
 Db GTGGGCTCCACCGCTGAGGCTCTTGACCTTATCGTTGAGCAATCGAGAGGCTGGCTTC 135638
 Qy 841 ACCCAGGCAAGGACATCGCTCTTGTCTCTGAGAGTTCCTCTCTGAGTTCTTCAAGGAC 900
 Db ACCCAGGCAAGGACATCGCTCTTGTCTCTGAGAGTTCCTCTCTGAGTTCTTCAAGGAC 135698
 Qy 901 GGCACTTACCACTTCGAAGGTGGCGACACTCCCGAGCTGAGATGGCAAACTTACGCT 960
 Db GGCACTTACCACTTCGAAGGTGGCGACACTCCCGAGCTGAGATGGCAAACTTACGCT 135758
 Qy 961 GAGCTCGTTGAGCGGTACCAATCGTCTCCATCGAGGACCCACTCGAGGAGATGACTGG 1020
 Db GAGCTCGTTGAGCGGTACCAATCGTCTCCATCGAGGACCCACTCGAGGAGATGACTGG 135818
 Qy 1021 GAGGTTTACACCAACCTCACCGCAACCATCGGCGACAAGTTTCAGATCGTTGGCGACGAC 1080
 Db GAGGTTTACACCAACCTCACCGCAACCATCGGCGACAAGTTTCAGATCGTTGGCGACGAC 135878
 Qy 1081 TTCTTCTGTCAACCACTTCAAGCGCTTGAAGAGGCGATTCGTAAAGAGGCTGCGCAACTCC 1140
 Db TTCTTCTGTCAACCACTTCAAGCGCTTGAAGAGGCGATTCGTAAAGAGGCTGCGCAACTCC 135938
 Qy 1141 ATCTGCTTGAAGTGAACCAAGTTCGTTACCTCACCGAGACCTTCGAGCGCTGTCGACATG 1200
 Db ATCTGCTTGAAGTGAACCAAGTTCGTTACCTCACCGAGACCTTCGAGCGCTGTCGACATG 135998
 Qy 1201 GCTCACCGCGAGGCTACACTCCATGATGTCCACCGTTCGCGTGAGACCGAGGACACC 1260
 Db GCTCACCGCGAGGCTACACTCCATGATGTCCACCGTTCGCGTGAGACCGAGGACACC 136058
 Qy 1261 ACCATTGCTGACCTCGAGTTGCACTCAACTGTGTGGCCAGATCAAGACTGTGTCTCCAGCA 1320
 Db ACCATTGCTGACCTCGAGTTGCACTCAACTGTGTGGCCAGATCAAGACTGTGTCTCCAGCA 136118
 Qy 1321 GGTTCGACCGTGTGCGCAAAAGTACAAACAGGTTCTCCGCATCGAGGAGCTGCTTGGCGAC 1380
 Db GGTTCGACCGTGTGCGCAAAAGTACAAACAGGTTCTCCGCATCGAGGAGCTGCTTGGCGAC 136178
 Qy 1381 GCGCGCTCTACGCGAGTCCGAGCGCATTCACCGCTTTCAGGGCTAAATAAAGCGCTT 1440
 Db GCGCGCTCTACGCGAGTCCGAGCGCATTCACCGCTTTCAGGGCTAAATAAAGCGCTT 136238
 Qy 1441 TTGACCGCCCGGTAACTCAAGGTTGCGGCGCTGTGTGCTTACTACTGTACTGTGCT 1500
 Db TTGACCGCCCGGTAACTCAAGGTTGCGGCGCTGTGTGCTTACTACTGTACTGTGCT 136298
 Qy 1501 GACTATGATCGAGATTTATGGCAAGCAGAGAAACTCATAAAGGCTTGTCTCTGTCT 1560
 Db GACTATGATCGAGATTTATGGCAAGCAGAGAAACTCATAAAGGCTTGTCTCTGTCT 136358
 Qy 1561 CAAGCAGGGAACGTGCTT 1578
 Db CAAGCAGGGAACGTGCTT 136376

RESULT 3

[illegible]

Qy 1191 TGTGACATGCTCACCGCGAGGCTACACCTTCATGATGTCCACCGTTCGGTGGAGAC 1250
 Db 1141 TGTGACATGCTCACCGCGAGGCTACACCTTCATGATGTCCACCGTTCGGTGGAGAC 1200
 Qy 1251 CGAGGACACCACTTGTGACCTCGAGTTGCACTCAACTGTGGCCAGATCAAGACTGG 1310
 Db 1201 CGAGGACACCACTTGTGACCTCGAGTTGCACTCAACTGTGGCCAGATCAAGACTGG 1260
 Qy 1311 TGCTCCAGCAGCTTCCGACCGTGTGCGAAAGTAAACACAGCTTCTCCGCGATCGAGCAGCT 1370
 Db 1261 TGCTCCAGCAGCTTCCGACCGTGTGCGAAAGTAAACACAGCTTCTCCGCGATCGAGCAGCT 1320
 Qy 1371 GCTTGGCGACCGCGCTTACGAGGTGCGAGGTGCGAGCGCATTCACCGCTTTCAGGCGTAAAT 1430
 Db 1321 GCTTGGCGACCGCGCTTACGAGGTGCGAGGTGCGAGCGCATTCACCGCTTTCAGGCGTAAAT 1380
 Qy 1431 AAAAGCGCTTTTCGACGCCCGGTAA 1455
 Db 1381 AAAAGCGCTTTTCGACGCCCGGTAA 1405

RESULT 4

ID AAF71395 standard; DNA; 1398 BP.

XX AAF71395;

XX AC

XX 30-APR-2001 (first entry)

DE Corynebacterium glutamicum SMP protein nucleotide sequence SEQ ID NO:71.

XX Corynebacterium glutamicum; carbon metabolism and energy production;
 KW SMP protein; sugar metabolism and oxidative phosphorylation protein;
 KW fine chemical production; organic acid; proteinogenic amino acid;
 KW nonproteinogenic amino acid; purine base; pyrimidine base; nucleoside;
 KW nucleotide; lipid; saturated fatty acid; unsaturated fatty acid; diol;
 KW carbohydrate; aromatic compound; vitamin; cofactor; polyketide; enzyme;
 KW diagnosis; Corynebacterium diphtheriae; evolutionary study; ds.

XX Corynebacterium glutamicum.

XX WO200100844-A2.

XX 04-JAN-2001.

XX 23-JUN-2000; 2000WO-IB000943.

XX 25-JUN-1999; 99US-0141031P.

PR 08-JUL-1999; 99DE-01031412.

PR 08-JUL-1999; 99DE-01031413.

PR 08-JUL-1999; 99DE-01031419.

PR 08-JUL-1999; 99DE-01031420.

PR 08-JUL-1999; 99DE-01031424.

PR 08-JUL-1999; 99DE-01031428.

PR 08-JUL-1999; 99DE-01031431.

PR 08-JUL-1999; 99DE-01031433.

PR 08-JUL-1999; 99DE-01031434.

PR 08-JUL-1999; 99DE-01031510.

PR 08-JUL-1999; 99DE-01031562.

PR 08-JUL-1999; 99DE-01031634.

PR 09-JUL-1999; 99DE-01032180.

PR 09-JUL-1999; 99DE-01032227.

PR 09-JUL-1999; 99DE-01032230.

PR 09-JUL-1999; 99US-0143208P.

PR 14-JUL-1999; 99DE-01032924.

PR 14-JUL-1999; 99DE-01032973.

PR 27-AUG-1999; 99DE-01040765.

PR 31-AUG-1999; 99US-0151572P.

PR 03-SEP-1999; 99DE-01042076.

PR 03-SEP-1999; 99DE-01042079.

PR 03-SEP-1999; 99DE-01042086.

PR 03-SEP-1999; 99DE-01042087.

PR 03-SEP-1999; 99DE-01042088.
 PR 03-SEP-1999; 99DE-01042095.
 PR 03-SEP-1999; 99DE-01042123.
 PR 03-SEP-1999; 99DE-01042125.
 XX (BADI) BASF AG.
 PA Pompejus M, Kroeger B, Schroeder H, Zelder O, Habehauer G;
 PI WPI; 2001-061975/07.
 DR P-PSDB; AAB79278.

XX New isolated Corynebacterium glutamicum nucleic acid encoding a sugar
 PT metabolism and oxidative phosphorylation protein for production or
 PT modulation of production of fine chemicals e.g. amino acids,
 PT carbohydrates or enzymes.

PS Claim 3; Page 242-244; 1246pp; English.

XX AAF71360 to AAF71750 encode the Corynebacterium glutamicum sugar
 CC metabolism and oxidative phosphorylation (SMP) proteins given in AAB79243
 CC to AAB 79633 which are involved in carbon metabolism and energy
 CC production. The C. glutamicum SMP gene can be used in vectors (II) for
 CC expression in host cells and production or modulation of production of
 CC fine chemicals, such as, an organic acid, a proteinogenic or
 CC nonproteinogenic amino acid (preferred), a purine or pyrimidine base, a
 CC nucleoside, a nucleotide, a lipid, a saturated or unsaturated fatty acid,
 CC a diol, a carbohydrate, an aromatic compound, a vitamin, a cofactor, a
 CC polyketide, or an enzyme. The presence of (I) or SMP proteins (III)
 CC encoded by them are used for diagnosing the presence or activity of
 CC Corynebacterium diphtheriae in a subject. (I), (II), (III) or host cells
 CC containing them are used to map genomes of organisms related to C.
 CC glutamicum, identify and localise C. glutamicum sequences of interest, in
 CC evolutionary studies, in determining SMP protein regions required for
 CC function, in modulating SMP protein activity, in modulating the
 CC metabolism of sugars, and in modulating high-energy molecule production
 CC in a cell (i.e. ATP, NADPH)

XX Sequence 1398 BP; 308 A; 423 C; 376 G; 291 T; 0 U; 0 Other;

Query Match 88.5%; Score 1396.4; DB 4; Length 1398;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1397; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 51 CGAAACAAGATTCTGTCACCAATTTGGGTGTAGACGTGATTGAAGACATTTTCATCACGTGA 110
 Db 1 CGAAACAAGATTCTGTCACCAATTTGGGTGTAGACGTGATTGAAGACATTTTCATCACGTGA 60
 Qy 111 ATAATCTAGTTAGTCTCCCAAGTTGGCATAGGAGGCCACAGTGGCTGAAATCATGCAGT 170
 Db 61 ATAATCTAGTTAGTCTCCCAAGTTGGCATAGGAGGCCACAGTGGCTGAAATCATGCAGT 120
 Qy 171 ATTGCTCGGCAATTTCTCGACTCCCGGGTAAACCAACCGTTCGAGGACAGGTTTTCCT 230
 Db 121 ATTGCTCGGCAATTTCTCGACTCCCGGGTAAACCAACCGTTCGAGGACAGGTTTTCCT 180
 Qy 231 GGATGACGGTTCCACCGGTGTCGAGGTGTTCCATCGGGCGCATCCACCGCGTCCACGA 290
 Db 181 GGATGACGGTTCCACCGGTGTCGAGGTGTTCCATCGGGCGCATCCACCGCGTCCACGA 240
 Qy 291 GGCTCATGAGTGGTGACGGTGGCGATCGCTACTCTGGCAAGGGCGTTTTTGAAGGCAGT 350
 Db 241 GGCTCATGAGTGGTGACGGTGGCGATCGCTACTCTGGCAAGGGCGTTTTTGAAGGCAGT 300
 Qy 351 TGAACACGTCAACGAAGAAATCGGCGACGAGCTCGCTGGGCTAGAGGCTGACGATCAGCG 410
 Db 301 TGAACACGTCAACGAAGAAATCGGCGACGAGCTCGCTGGGCTAGAGGCTGACGATCAGCG 360
 Qy 411 CCTCATCGACGACGATGATCAGCTTGATGGCACCGCCCAACAAAGTCCCGCTGGGTGC 470
 Db 361 CCTCATCGACGACGATGATCAGCTTGATGGCACCGCCCAACAAAGTCCCGCTGGGTGC 420
 Qy 471 AACGCAATCCTTGGTGTTCATCGGCTGTTTGCACAAAGGCTGCTGCTGATTCGCGAGGCCT 530

Db	421		AAA	CGCAATCTTTGGTGTTCATAGCTGTTCAAAGGCTGCTGTGATTTCCGCAAGGCT	480
Qy	531	CCCACTGTTCCGCTACATCGGTGGACAAACGACACACGTTCTTCCAGTTCCCAATGATGAA	590		
Db	481	CCCACTGTTCCGCTACATCGGTGGACAAACGACACGTTCTTCCAGTTCCCAATGATGAA	540		
Qy	591	CATCATCAACGGTGGGCTCACGCTGACTCCGGTGTGTGACGTTGAGAAATTCATGATCGC	650		
Db	541	CATCATCACCGTGGGCTCACGCTGACTCCGGTGTGTGACGTTGAGAAATTCATGATCGC	600		
Qy	651	TCCAAATCGGTGACAGACCTTCTCTGAGGCTCTCCGAAACGGCGCGAGGTCTACACGC	710		
Db	601	TCCAAATCGGTGACAGACCTTCTCTGAGGCTCTCCGAAACGGCGCGAGGTCTACACGC	660		
Qy	711	ACTGAAGTCGCTCATCAAGGAAAAGGCGCTGTCCACCGGACTTGGCGATGAGGGGCGCTT	770		
Db	661	ACTGAAGTCGCTCATCAAGGAAAAGGCGCTGTCCACCGGACTTGGCGATGAGGGGCGCTT	720		
Qy	771	CGCTCTCTCGTCGGCTCCACCGTGGAGGCTCTTGACCTTATCGTTGAGCGAATCGAGAA	830		
Db	721	CGCTCTCTCGTCGGCTCCACCGTGGAGGCTCTTGACCTTATCGTTGAGCGAATCGAGAA	780		
Qy	831	GGCTGGCTTCACCCCAGGCAAGACATCGCTCTTGCTCTGGAAGTTGCTTCTCTCAGTT	890		
Db	781	GGCTGGCTTCACCCCAGGCAAGACATCGCTCTTGCTCTGGAAGTTGCTTCTCTCAGTT	840		
Qy	891	CTTCAAGGACGGCACCTACACATTCGAAAGTGGCGAGCACTCCGCAAGCTGAGATGCCAAA	950		
Db	841	CTTCAAGGACGGCACCTACACATTCGAAAGTGGCGAGCACTCCGCAAGCTGAGATGCCAAA	900		
Qy	951	CGTTTACGCTGAGCTCGTTGACGGGTACCCAAATCGTCTCCATCGAGACCCACTGTCAGGA	1010		
Db	901	CGTTTACGCTGAGCTCGTTGACGGGTACCCAAATCGTCTCCATCGAGACCCACTGTCAGGA	960		
Qy	1011	AGATGACTGGGAGGGTTACACCAACTCAACGGCAACATCGGCGACAAGGTTCCAGATCGT	1070		
Db	961	AGATGACTGGGAGGGTTACACCAACTCAACGGCAACATCGGCGACAAGGTTCCAGATCGT	1020		
Qy	1071	TGGCGAGCACTCTTCGCTCAACCACTGAGCGCTGAAAGAGGGATCGCTAAGAAGGC	1130		
Db	1021	TGGCGAGCACTCTTCGCTCAACCACTGAGCGCTGAAAGAGGGATCGCTAAGAAGGC	1080		
Qy	1131	TGCCAACTCCATCTCGTTAAGGTGAACCCAGATCGGTACCTCACCGAGACCTTCGACGC	1190		
Db	1081	TGCCAACTCCATCTCGTTAAGGTGAACCCAGATCGGTACCTCACCGAGACCTTCGACGC	1140		
Qy	1191	TGTCGACATGGCTCACCGCGAGGTACCTCCATGATCTCCACCGTTCCGGTCGAGAC	1250		
Db	1141	TGTCGACATGGCTCACCGCGAGGTACCTCCATGATCTCCACCGTTCCGGTCGAGAC	1200		
Qy	1251	CGAGGACACCACTTGCTGACCTTCGAGTTGCACTCAACTGTGGCCAGATCAAGACTGG	1310		
Db	1201	CGAGGACACCACTTGCTGACCTTCGAGTTGCACTCAACTGTGGCCAGATCAAGACTGG	1260		
Qy	1311	TGCTCCAGACGTTCCGACCGGTGCGAAAGTACAACACAGCTTCTCCGATCGAGCAGCT	1370		
Db	1261	TGCTCCAGACGTTCCGACCGGTGCGAAAGTACAACACAGCTTCTCCGATCGAGCAGCT	1320		
Qy	1371	GCTTGGGACGCGCGGCTACTACGAGGTTCGAGGGCATTCGCCAGCTTTCAGGGCTAAAT	1430		
Db	1321	GCTTGGGACGCGCGGCTACTACGAGGTTCGAGGGCATTCGCCAGCTTTCAGGGCTAAAT	1380		
Qy	1431	AAAAGCGCTTTTCGACGC	1448		
Db	1381	AAAAGCGCTTTTCGACGC	1398		

RESULT 5
ACC45311
ID ACC45311 standard; DNA; 1278 BP.
XX
AC ACC45311;

XX		
DT	17-JUN-2003	(first entry)
DE	Corynebacterium glutamicum	eno encoding DNA SEQ ID NO:1.
XX		
KW	Fine chemical; Coryneform bacteria;	Escherichia coli; microorganism;
KW	genetically modified microorganism;	metabolite; biosynthesis; amino acid;
KW	vitamin; nucleoside; nucleotide;	pigment; protein; human medicine;
KW	pharmaceutical; food; animal feeding;	eno; gene; ds.
XX		
OS	Corynebacterium glutamicum.	
XX		
PH	Key	Location/Qualifiers
FT	CDS	1..1278
FT		/*tag= a
FT		/product= "eno protein"
XX		
PN	WO2003023016-A2.	
XX		
PD	20-MAR-2003.	
XX		
XX	11-SEP-2002; 2002WO-EP010174.	
Pf		
PR	13-SEP-2001; 2001DE-01045043.	
XX		
PA	(DEGS) DEGUSSA AG.	
XX		
PI	Farwick M, Hermann T;	
XX		
DR	WPI; 2003-354534/33.	
DR	P-PsDB; ABP97001.	
XX		
PT	Microorganism useful for producing e.g. fine chemicals,	has permanently altered phosphorylatability protein, such that biosynthesis of fine chemical synthesized by microorganism is increased compared to wild-type.
PT		
XX		
PS	Example 3; Page 78-80; 120pp; English.	
XX		
CC	The present invention describes a microorganism (I), in which the phosphorylatability of at least one protein has been permanently altered such that the biosynthesis of at least one fine chemical synthesised by the microorganism is increased compared to the wild type. Also described:	
CC	(1) use of a DNA (II) sequence coding for a protein which contains a phosphorylation site, where the sequence contains such a mutation that the protein is changed in its phosphorylatability for the production of (I), or for the production of fine chemicals; and (2) a method for producing fine chemicals or metabolites comprising using (I). (I) is useful for producing fine chemicals or metabolites, such as amino acids, vitamins, nucleosides, nucleotides, pigments or proteins. The amino acids and vitamins produced using (I) can be used in human medicine, in the pharmaceutical industry, food industry and in animal feeding. (I) produces larger amount of desired fine chemical or a metabolite than the wild type. The present sequence encodes wild type eno from Corynebacterium glutamicum, which is used in an example from the present invention	
CC		
XX	Sequence 1278 BP; 271 A; 399 C; 347 G; 261 T; 0 U; 0 Other;	

Qy 331 AAGGCGGTTTGAAGGCGAGTTGAAAACTCAACGAAGAAATCGGCGACGAGCTCGCTGGC 390
 Db |||||
 Qy 181 AAGGCGGTTTGAAGGCGAGTTGAAAACTCAACGAAGAAATCGGCGACGAGCTCGCTGGC 240
 Db |||||
 Qy 391 CTAGAGGCTGACGATCAGCGGCTCATGCAAGCAATGATCAAGCTTGATGGCAACGCC 450
 Db |||||
 Qy 241 CTAGAGGCTGACGATCAGCGGCTCATGCAAGCAATGATCAAGCTTGATGGCAACGCC 300
 Db |||||
 Qy 451 AACAAAGTCCCGCTGGGTGCAAAACGCAATCTTGGTGTTCATGGCTGTGCAAAAGCT 510
 Db |||||
 Qy 301 AACAAAGTCCCGCTGGGTGCAAAACGCAATCTTGGTGTTCATGGCTGTGCAAAAGCT 360
 Db |||||
 Qy 511 GCTGCTGATTCCGAGGCTCCCACTGTTCCGCTACATCGGTGACCAAAACGACACGTT 570
 Db |||||
 Qy 361 GCTGCTGATTCCGAGGCTCCCACTGTTCCGCTACATCGGTGACCAAAACGACACGTT 420
 Db |||||
 Qy 571 CTTCCAGTCCCAATGATGAACATCATCAACGGTGGCGCTACGCTGACTCCGGTGTGAC 630
 Db |||||
 Qy 421 CTTCCAGTCCCAATGATGAACATCATCAACGGTGGCGCTACGCTGACTCCGGTGTGAC 480
 Db |||||
 Qy 631 GTTCAGGAATTCATGATCGCTCCCAATCGGTGACAGACCTTCTCTGAGGCTCTCCGCAAC 690
 Db |||||
 Qy 481 GTTCAGGAATTCATGATCGCTCCCAATCGGTGACAGACCTTCTCTGAGGCTCTCCGCAAC 540
 Db |||||
 Qy 691 GGC CGGAGGTCTACCAAGCACTGAAGTCCGTCATCAAGGAAAGGCGCTGTCCACCGGA 750
 Db |||||
 Qy 541 GGC CGGAGGTCTACCAAGCACTGAAGTCCGTCATCAAGGAAAGGCGCTGTCCACCGGA 600
 Db |||||
 Qy 751 CTTGGCGATGAGGCGGCTTCCGCTTCCGCTCGGCTCCACCCGTGAGGCTCTTGACCTT 810
 Db |||||
 Qy 601 CTTGGCGATGAGGCGGCTTCCGCTTCCGCTCGGCTCCACCCGTGAGGCTCTTGACCTT 660
 Db |||||
 Qy 811 ATCGTTGAGCAATCGAGAGGCTGCTTCAACCCAGGAGGATCGCTCTGCTGCTG 870
 Db |||||
 Qy 661 ATCGTTGAGCAATCGAGAGGCTGCTTCAACCCAGGAGGATCGCTCTGCTGCTG 720
 Db |||||
 Qy 871 GAGCTTCTCTCTGAGTCTTCAAGGACGGCACTTACCACCTTCAAGAGTGGCGCAGCAC 930
 Db |||||
 Qy 721 GAGCTTCTCTCTGAGTCTTCAAGGACGGCACTTACCACCTTCAAGAGTGGCGCAGCAC 780
 Db |||||
 Qy 931 TCCGACGCTGAGTGGCAAGCTTTAGCTGAGCTCGTTGACGCGTACCAATCGTCTCC 990
 Db |||||
 Qy 781 TCCGACGCTGAGTGGCAAGCTTTAGCTGAGCTCGTTGACGCGTACCAATCGTCTCC 840
 Db |||||
 Qy 991 ATCGAGGACCCACTGCGAGGAGATGACTGGAGGTTTACACCACTTCAACCAACATC 1050
 Db |||||
 Qy 841 ATCGAGGACCCACTGCGAGGAGATGACTGGAGGTTTACACCACTTCAACCAACATC 900
 Db |||||
 Qy 1051 GGCACAAAGTTTCAGATCGTTGGCGAGCACTTCTTCGTCACCAACCTTGAGCGCTGAAG 1110
 Db |||||
 Qy 901 GGCACAAAGTTTCAGATCGTTGGCGAGCACTTCTTCGTCACCAACCTTGAGCGCTGAAG 960
 Db |||||
 Qy 1111 GAGGCGATCGTAAGAGGCTGCCAATCCATCTCGTTTAAAGTGAACAGATCGGTACC 1170
 Db |||||
 Qy 961 GAGGCGATCGTAAGAGGCTGCCAATCCATCTCGTTTAAAGTGAACAGATCGGTACC 1020
 Db |||||
 Qy 1171 CTCACCGAGACCTTCGAGCGTGTGACATGCTTCAACCGCGAGGCTACCTCCATGATG 1230
 Db |||||
 Qy 1021 CTCACCGAGACCTTCGAGCGTGTGACATGCTTCAACCGCGAGGCTACCTCCATGATG 1080
 Db |||||
 Qy 1231 TCCACCGCTTCCGCTGAGACCGAGGACCAACCAATTCGTACCTTCGAGTTGCACTCAAC 1290
 Db |||||
 Qy 1081 TCCACCGCTTCCGCTGAGACCGAGGACCAACCAATTCGTACCTTCGAGTTGCACTCAAC 1140
 Db |||||
 Qy 1291 TGTGGCGAGATCAAGACTGTTGCTCCAGCAGGTTCCGACCGTGTCCGAAAGTACCAACAG 1350
 Db |||||
 Qy 1141 TGTGGCGAGATCAAGACTGTTGCTCCAGCAGGTTCCGACCGTGTCCGAAAGTACCAACAG 1200
 Db |||||
 Qy 1351 CTTCTCCGATCGAGGAGCTGCTTGGCGAGCGCGGCTTACCGAGGTCGAGGCGGATTC 1410
 Db |||||
 Qy 1201 CTTCTCCGATCGAGGAGCTGCTTGGCGAGCGCGGCTTACCGAGGTCGAGGCGGATTC 1260
 Db |||||

Qy 1411 CCACGCTTTTCAGGGCTAA 1428
 Db |||||
 Qy 1261 CCACGCTTTTCAGGGCTAA 1278
 Db |||||

RESULT 6

ID AAH66050
 AAH66050 standard; DNA; 1275 BP.

XX AC AAH66050;

XX 26-SEP-2001 (first entry)

DE C glutamicum coding sequence fragment SEQ ID NO: 1085.

XX Corynebacterium; amino acid synthesis; vitamin; saccharide;

XX organic acid synthesis; ds.

XX Corynebacterium glutamicum.

PN EP1108790-A2.

PD 20-JUN-2001.

XX 18-DEC-2000; 2000EP-00127688.

XX 16-DEC-1999; 99JP-00377484.

PR 07-APR-2000; 2000JP-00159162.

PR 03-AUG-2000; 2000JP-00280988.

XX (KYOW) KYOWA HAKKO KOGYO KK.

XX Nakagawa S, Mizouchi H, Ando S, Hayaashi M, Ochiai K, Yokoi H;

PI Tateishi N, Senoh A, Ikeda M, Ozaki A;

XX WPI; 2001-376931/40.

DR P-PSDE; AAG90831.

XX Novel polynucleotides derived from Corynebacterium bacteria, for identifying

PT mutation point of a gene, measuring expression of a gene, analyzing

PT expression profile or pattern of a gene and identifying homologous gene.

XX Claim 8; SEQ ID NO 1085; 246pp + Sequence Listing; English.

XX The present invention provides a number of nucleotide and protein

CC sequences from the Corynebacterium bacterium Corynebacterium glutamicum. These

CC are useful for identifying the mutation point of a gene derived from a

CC mutant of corynebacterium bacterium, measuring expression amount and analysing

CC the expression profile or expression pattern of a gene derived from

CC Corynebacterium bacterium, and identifying a homologue of a gene derived from

CC Corynebacterium bacterium. Corynebacterium bacteria are useful for producing amino

CC acids, nucleic acids, vitamins, saccharides and organic acids,

CC particularly L-lysine. The present sequence is a nucleic acid described

CC in the exemplification of the invention. Note: The sequence data for this

CC patent did not form part of the printed specification, but was obtained

CC in electronic format directly from the European Patent Office

XX

SQ Sequence 1275 BP; 269 A; 399 C; 347 G; 260 T; 0 U; 0 Other;

Query Match 80.8%; Score 1275; DB 5; Length 1275;

Best Local Similarity 100.0%; Pred. No. 3.6e-311;

Matches 1275; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 151 GTGGCTGAATCATGACGCTATTTCGCTCGGAAATTCGCACTCCCGGGTAACCCAAAC 210

Db 1 GTGGCTGAATCATGACGCTATTTCGCTCGGAAATTCGCACTCCCGGGTAACCCAAAC 60

Qy 211 GTGAGGCGAGAGGTTTTCTTGGATGACGTTTCCACGCTGTCGAGGTGTTCCATCCGGC 270

Db 61 GTGAGGCGAGAGGTTTTCTTGGATGACGTTTCCACGCTGTCGAGGTGTTCCATCCGGC 120

Qy 271 GCATCCACCGCGCTCCACGAGGCTCATGAGCTCGTGACGCTGGCGATCGCTACCTCGGC 330

Db |||||

Db 121 GCATCCACCGCGCTCCACGAGGCTCATGAGTCGCTGACGGTGGCGATCGCTACCTGGGC 180
 Qy 331 AAGGCGCTTTTGAAGGCGAGTTGAAACGTCGTAAGAGAAATCGCGAGCGAGCTCGCTGGC 390
 Db 181 AAGGCGCTTTTGAAGGCGAGTTGAAACGTCGTAAGAGAAATCGCGAGCGAGCTCGCTGGC 240
 Qy 391 CTAGAGGCTCAGCATCAGCGCTCATCGACGAGCAATGATCAAGCTTGATGCGACCGCC 450
 Db 241 CTAGAGGCTCAGCATCAGCGCTCATCGACGAGCAATGATCAAGCTTGATGCGACCGCC 300
 Qy 451 AACAAAGTCCCGCTGGGTGCAACGCAATCTTGGTGTTCATGGCTGTGTGCAAGGCT 510
 Db 301 AACAAAGTCCCGCTGGGTGCAACGCAATCTTGGTGTTCATGGCTGTGTGCAAGGCT 360
 Qy 511 GCTGCTGATTCGCGAGGCTCCCACTGCTTCGCTACATCGGTGAGCAAAAGCACAGTT 570
 Db 361 GCTGCTGATTCGCGAGGCTCCCACTGCTTCGCTACATCGGTGAGCAAAAGCACAGTT 420
 Qy 571 CTTCCAGTTCGAATGATGAACATCAACGCGTGGCTCAGCGTGAATCCGGTGTGAC 630
 Db 421 CTTCCAGTTCGAATGATGAACATCAACGCGTGGCTCAGCGTGAATCCGGTGTGAC 480
 Qy 631 GTTCAGGAATTCATGATCGCTCCAAATCGGTGACAGACCTTCTCTGAGGCTCTCCGCAAC 690
 Db 481 GTTCAGGAATTCATGATCGCTCCAAATCGGTGACAGACCTTCTCTGAGGCTCTCCGCAAC 540
 Qy 691 GGCGCGAGGTCTACCGACGATGAAGTCCGTGATCAAGGAAAGGGCCTGTCCACCGGA 750
 Db 541 GGCGCGAGGTCTACCGACGATGAAGTCCGTGATCAAGGAAAGGGCCTGTCCACCGGA 600
 Qy 751 CTTGCGCATGAGGCGGCTTCGCTCCTTCGCTGGCTCCACCGTGAAGGCTCTTGACCTT 810
 Db 601 CTTGCGCATGAGGCGGCTTCGCTCCTTCGCTGGCTCCACCGTGAAGGCTCTTGACCTT 660
 Qy 811 ATCGTTGAGGCAATCGAGAAGGCTGGCTTCAACCGAGGAGGATCGCTCTTGCTCTG 870
 Db 661 ATCGTTGAGGCAATCGAGAAGGCTGGCTTCAACCGAGGAGGATCGCTCTTGCTCTG 720
 Qy 871 GACGTTGCTTCTGAGTCTTCAAGGACGGACCTTACGATTCGAGGTGGCGACGAC 930
 Db 721 GACGTTGCTTCTGAGTCTTCAAGGACGGACCTTACGATTCGAGGTGGCGACGAC 780
 Qy 931 TCCGACGCTGAGTGGCAAGCTTTACGCTGAGCTGCTGAGCGGTACCCAAATCGTCTCC 990
 Db 781 TCCGACGCTGAGTGGCAAGCTTTACGCTGAGCTGCTGAGCGGTACCCAAATCGTCTCC 840
 Qy 991 ATCGAGGACCCACTGCAGGAGATGACCTGGAGGGTTACACCAACCTCACCGCAACCATC 1050
 Db 841 ATCGAGGACCCACTGCAGGAGATGACCTGGAGGGTTACACCAACCTCACCGCAACCATC 900
 Qy 1051 GCGGACAAAGTTGAGTGGCGAGCGAATCTTCTGTCACCAACCTGAGCGCTGAAG 1110
 Db 901 GCGGACAAAGTTGAGTGGCGAGCGAATCTTCTGTCACCAACCTGAGCGCTGAAG 960
 Qy 1111 GAGGCGATCGTAAAGGCTGCCAATCTCCTGTTAGGTGAACGATCGGTACC 1170
 Db 961 GAGGCGATCGTAAAGGCTGCCAATCTCCTGTTAGGTGAACGATCGGTACC 1020
 Qy 1171 CTCACCGAGACCTTCGACGCTGTGCACATGCTCACCGCGAGGCTACACCTCCATGATG 1230
 Db 1021 CTCACCGAGACCTTCGACGCTGTGCACATGCTCACCGCGAGGCTACACCTCCATGATG 1080
 Qy 1231 TCCACCGTTCGCTGAGACCGAGGACACCACTTGTGACCTCGAGTTGCACTCAAC 1290
 Db 1081 TCCACCGTTCGCTGAGACCGAGGACACCACTTGTGACCTCGAGTTGCACTCAAC 1140
 Qy 1291 TGTGCGCAGATCAAGACTGCTGCTCCAGCAGTTCGACCGGTGTCGCAAAAGTCAACACAG 1350
 Db 1141 TGTGCGCAGATCAAGACTGCTGCTCCAGCAGTTCGACCGGTGTCGCAAAAGTCAACACAG 1200
 Qy 1351 CTTCTCCGATCGAGCAGCTCTTGGCGACCGCGGCTTACGAGGTTCGAGCGCATTC 1410
 Db 1201 CTTCTCCGATCGAGCAGCTCTTGGCGACCGCGGCTTACGAGGTTCGAGCGCATTC 1260

Qy 1411 CCACGCTTTTCAGGC 1425
 Db 1261 CCACGCTTTTCAGGC 1275

RESULT 7

ACC45312
 ID ACC45312 standard; DNA; 1299 BP.

XX ACC45312;

DT 17-JUN-2003 (first entry)

DE Corynebacterium glutamicum eno DNA fragment SEQ ID NO:3.

XX Fine chemical; Corynebacterium bacteria; Escherichia coli; microorganism;
 KW genetically modified microorganism; metabolite; biosynthesis; amino acid;
 KW vitamin; nucleoside; nucleotide; pigment; protein; human medicine;
 KW pharmaceutical; food; animal feeding; eno; gene; ds.

OS Corynebacterium glutamicum.

XX WO2003023016-A2.

PN 20-MAR-2003.

XX 11-SEP-2002; 2002WO-EP010174.

XX 13-SEP-2001; 2001DB-01045043.

XX (DEGS) DEGUSSA AG.

XX Farwick M, Hermann T;

XX WPI; 2003-354534/33.

Microorganism useful for producing e.g. fine chemicals, has permanently altered phosphorylatability protein, such that biosynthesis of fine chemical synthesized by microorganism is increased compared to wild-type.

Example 3; Page 81-82; 120pp; English.

The present invention describes a microorganism (I), in which the phosphorylatability of at least one protein has been permanently altered such that the biosynthesis of at least one fine chemical synthesised by the microorganism is increased compared to the wild type. Also described: (1) use of a DNA (II) sequence coding for a protein which contains a phosphorylation site, where the sequence contains such a mutation that the protein is changed in its phosphorylatability for the production of (1), or for the production of fine chemicals; and (2) a method for producing fine chemicals or metabolites comprising using (I). (I) is useful for producing fine chemicals or metabolites, such as amino acids, vitamins, nucleosides, nucleotides, pigments or proteins. The amino acids and vitamins produced using (I) can be used in human medicine, in the pharmaceutical industry, food industry and in animal feeding. (I) produces larger amount of desired fine chemical or a metabolite than the wild type. The present sequence represents a wild type eno fragment DNA sequence from Corynebacterium glutamicum, which is used in an example from the present invention

Sequence 1299 BP; 274 A; 409 C; 353 G; 263 T; 0 U; 0 Other;

Query Match 80.8%; Score 1275; DB 8; Length 1299;
 Best Local Similarity 100.0%; Pred. No. 3.6e-311;
 Matches 1275; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 154 GCTGAATCATGACGATTTTCGCTCGGAAATTCGACTCCCGCGTAAACCCACCGTC 213
 Db 10 GCTGAATCATGACGATTTTCGCTCGGAAATTCGACTCCCGCGTAAACCCACCGTC 69
 Qy 214 GAGCAGAGGTTTTCTTGGATGACGGTCCACGGTTCGAGGTTCATCCGCGGCA 273

Db 70 GAGGAGAGGTTTCTCGATGACGGTTCCACCGGTGTCGAGGTGTTCCATCCGGGCA 129
 Qy 274 TCCACCGCGCTCCACAGGCTCATGAGCTGCGTGAACGGTGGCGATCGCTACCTGGGCAAG 333
 Db 130 TCCACCGCGCTCCACAGGCTCATGAGCTGCGTGAACGGTGGCGATCGCTACCTGGGCAAG 189
 Qy 334 GCGCTTTTGAAGGAGTTGAAAAAGTCAACGAAAGAAATCGCGCGACGAGCTCGCTGGGCTA 393
 Db 190 GCGCTTTTGAAGGAGTTGAAAAAGTCAACGAAAGAAATCGCGCGACGAGCTCGCTGGGCTA 249
 Qy 394 GAGCTGACGATCAGCGCTCATCGAAGCAATGATCAAGCTTGATGGCAACCGCCAAC 453
 Db 250 GAGCTGACGATCAGCGCTCATCGAAGCAATGATCAAGCTTGATGGCAACCGCCAAC 309
 Qy 454 AAGTCCCGCTGGGTGCAACGCAATCTTGGTGTTCATGGCTGTGTCAAAGGCTGCT 513
 Db 310 AAGTCCCGCTGGGTGCAACGCAATCTTGGTGTTCATGGCTGTGTCAAAGGCTGCT 369
 Qy 514 GCTGATTCGGGAGGCTCCGACTGTTCCGCTACATCGGTGGAACCAAGCGACACGTTCTT 573
 Db 370 GCTGATTCGGGAGGCTCCGACTGTTCCGCTACATCGGTGGAACCAAGCGACACGTTCTT 429
 Qy 574 CCAGTCCCAATGATGAACATCATCAACGGTGGCGCTCAGCTGACTCGGTGTTGACGTT 633
 Db 430 CCAGTCCCAATGATGAACATCATCAACGGTGGCGCTCAGCTGACTCGGTGTTGACGTT 489
 Qy 634 CAGGAATTCATGATCGCTCCAAATCGGTGCGAGACCTTCTCTGAGGCTCTCCGCAACGGC 693
 Db 490 CAGGAATTCATGATCGCTCCAAATCGGTGCGAGACCTTCTCTGAGGCTCTCCGCAACGGC 549
 Qy 694 GCGAGGTCTACCAAGCAATGAAATCGGTGCAATCAAGGAAAGGCGCTGTCCACCGGACTT 753
 Db 550 GCGAGGTCTACCAAGCAATGAAATCGGTGCAATCAAGGAAAGGCGCTGTCCACCGGACTT 609
 Qy 754 GCGATGAGGCGGCTGCTGCTCTCCGCTCGGTGCGCTCCACCGTGGGCTCTTGACCTTATC 813
 Db 610 GCGATGAGGCGGCTGCTGCTCTCCGCTCGGTGCGCTCCACCGTGGGCTCTTGACCTTATC 669
 Qy 814 GTTGAGGCAATCGAAGAGGCTGGCTTCAACCCAGGCAAGGACATCGCTCTTGCTCTGGAC 873
 Db 670 GTTGAGGCAATCGAAGAGGCTGGCTTCAACCCAGGCAAGGACATCGCTCTTGCTCTGGAC 729
 Qy 874 GTTGCTTCTCTGAGTTCTTCAAGGAGCGCACTTACCACTTCGAAGGTGGCCAGCACTCC 933
 Db 730 GTTGCTTCTCTGAGTTCTTCAAGGAGCGCACTTACCACTTCGAAGGTGGCCAGCACTCC 789
 Qy 934 GCAGCTGAGATGGCAAAAGTTTACGCTGAGCTGTTGACCGGTACCACTTCGCTCCATC 993
 Db 790 GCAGCTGAGATGGCAAAAGTTTACGCTGAGCTGTTGACCGGTACCACTTCGCTCCATC 849
 Qy 994 GAGGACCACTTCGAGGAGATGACTGGGAGGTTTACACCACTTCACCGCAACCACTCGC 1053
 Db 850 GAGGACCACTTCGAGGAGATGACTGGGAGGTTTACACCACTTCACCGCAACCACTCGC 909
 Qy 1054 GACAAGTTTCAGATCGTTGGCGAGCTTCTTCTGTCACCACTTCGAGCGCTGAAGGAG 1113
 Db 910 GACAAGTTTCAGATCGTTGGCGAGCTTCTTCTGTCACCACTTCGAGCGCTGAAGGAG 969
 Qy 1114 GGCATCGCTAAGAGGCTGCCAATCCATCTGTTAAGGTGAACGAGATCGGTACCCCTC 1173
 Db 970 GGCATCGCTAAGAGGCTGCCAATCCATCTGTTAAGGTGAACGAGATCGGTACCCCTC 1029
 Qy 1174 ACCGAGACCTTCGAGCTGTGACATGCTCACCAGCGAGCTACCTCATGATGCTC 1233
 Db 1030 ACCGAGACCTTCGAGCTGTGACATGCTCACCAGCGAGCTACCTCATGATGCTC 1089
 Qy 1234 CACCGTTCCGGTGAGACCGGAGGACCAACCATTCCTGACCTCGAGTTGCACTCAACTGT 1293
 Db 1090 CACCGTTCCGGTGAGACCGGAGGACCAACCATTCCTGACCTCGAGTTGCACTCAACTGT 1149
 Qy 1294 GGCAGATCAAGACTGTGCTCCAGCAGCTTCCGACCGGTGTCGCAAAAGTCAACACGCTT 1353
 Db 1150 GGCAGATCAAGACTGTGCTCCAGCAGCTTCCGACCGGTGTCGCAAAAGTCAACACGCTT 1209

Qy 1354 CTCCGCATCGAGCAGCTGCTTGGCGACGCCGGCTCTACGAGGTGCGAGCGCATTTCCA 1413
 Db 1210 CTCCGCATCGAGCAGCTGCTTGGCGACGCCGGCTCTACGAGGTGCGAGCGCATTTCCA 1269
 Qy 1414 CGCTTTCAGGGCTAA 1428
 Db 1270 CGCTTTCAGGGCTAA 1284
 RESULT 8
 ACC45314
 ID ACC45314 standard; DNA; 1278 BP.
 XX ACC45314;
 DX 17-JUN-2003 (first entry)
 XX Modified eno S330E encoding DNA SEQ ID NO:5.
 DE
 XX
 XX
 KW genetically modified microorganism; Escherichia coli; microorganism;
 KW vitamin; nucleoside; nucleotide; pigment; protein; human medicine;
 KW pharmaceutical; food; animal feeding; eno; gene; ds.
 XX Corynebacterium glutamicum.
 OS
 XX Synthetic.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..1278
 FT /*tag= a
 FT /product= "modified eno S330E protein"
 PN
 XX W02003023016-A2.
 PD
 XX 20-MAR-2003.
 PF 11-SEP-2002; 2002WO-EP010174.
 PR 13-SEP-2001; 2001DE-01045043.
 XX (DEGS) DEGUSSA AG.
 PA Farwick M, Hermann T;
 PI WPI; 2003-354534/33.
 DR P-PSDB; ABP97002.
 XX
 PT Microorganism useful for producing e.g. fine chemicals, has permanently
 PT altered phosphorylatability protein, such that biosynthesis of fine
 PT chemical synthesized by microorganism is increased compared to wild-type.
 XX
 PS Example 3; Page 83-84; 120pp; English.
 XX
 CC The present invention describes a microorganism (I), in which the
 CC phosphorylatability of at least one protein has been permanently altered
 CC such that the biosynthesis of at least one fine chemical synthesised by
 CC the microorganism is increased compared to the wild type. Also described:
 CC (1) use of a DNA (II) sequence coding for a protein which contains a
 CC phosphorylation site, where the sequence contains such a mutation that
 CC the protein is changed in its phosphorylatability for the production of
 CC (1), or for the production of fine chemicals; and (2) a method for
 CC producing fine chemicals or metabolites comprising using (I). (I) is
 CC useful for producing fine chemicals or metabolites, such as amino acids,
 CC vitamins, nucleosides, nucleotides, pigments or proteins. The amino acids
 CC and vitamins produced using (I) can be used in human medicine, in the
 CC pharmaceutical industry, food industry and in animal feeding. (I)
 CC produces larger amount of desired fine chemical or a metabolite than the
 CC wild type. The present sequence encodes a modified eno S330E from
 CC Corynebacterium glutamicum, which is used in an example from the present
 CC invention
 XX
 XX Sequence 1278 BP; 272 A; 397 C; 349 G; 260 T; 0 U; 0 Other;

Query Match 80.7%; Score 1273.2; DB 8; Length 1278; Best Local Similarity 99.8%; Pred. No. 1e-310; Matches 1275; Conservative 0; Mismatches 3; Indels 0; Gaps 0;									
QY	151	GTGGCTGAATCATGACGATTTTCGCTCGGAAATTC	CGGAAATTC	CGGAAATTC	CGGAAATTC	CGGAAATTC	CGGAAATTC	CGGAAATTC	CGGAAATTC
DB	1	GTGGCTGAATCATGACGATTTTCGCTCGGAAATTC	CGGAAATTC	CGGAAATTC	CGGAAATTC	CGGAAATTC	CGGAAATTC	CGGAAATTC	CGGAAATTC
QY	211	GTGAGGACAGAGTTTTCCTGGATGACGGTTC	CGGAAATTC	CGGAAATTC	CGGAAATTC	CGGAAATTC	CGGAAATTC	CGGAAATTC	CGGAAATTC
DB	61	GTGAGGACAGAGTTTTCCTGGATGACGGTTC	CGGAAATTC	CGGAAATTC	CGGAAATTC	CGGAAATTC	CGGAAATTC	CGGAAATTC	CGGAAATTC
QY	271	GCATCCACCGGCTCCACGAGGCTCATGAGCT	CGGAAATTC	CGGAAATTC	CGGAAATTC	CGGAAATTC	CGGAAATTC	CGGAAATTC	CGGAAATTC
DB	121	GCATCCACCGGCTCCACGAGGCTCATGAGCT	CGGAAATTC	CGGAAATTC	CGGAAATTC	CGGAAATTC	CGGAAATTC	CGGAAATTC	CGGAAATTC
QY	331	AAGGCGTTTTGAAGGCTGAAACGCTCAACG	AGAAATTC	AGAAATTC	AGAAATTC	AGAAATTC	AGAAATTC	AGAAATTC	AGAAATTC
DB	181	AAGGCGTTTTGAAGGCTGAAACGCTCAACG	AGAAATTC	AGAAATTC	AGAAATTC	AGAAATTC	AGAAATTC	AGAAATTC	AGAAATTC
QY	391	CTAGAGCTGACGATCAGCGCTCATCGACGA	AGAAATTC	AGAAATTC	AGAAATTC	AGAAATTC	AGAAATTC	AGAAATTC	AGAAATTC
DB	241	CTAGAGCTGACGATCAGCGCTCATCGACGA	AGAAATTC	AGAAATTC	AGAAATTC	AGAAATTC	AGAAATTC	AGAAATTC	AGAAATTC
QY	451	AACAGTCCCGCTGGGTGMAACGCAATCTTG	GGTGTTC	GGTGTTC	GGTGTTC	GGTGTTC	GGTGTTC	GGTGTTC	GGTGTTC
DB	301	AACAGTCCCGCTGGGTGMAACGCAATCTTG	GGTGTTC	GGTGTTC	GGTGTTC	GGTGTTC	GGTGTTC	GGTGTTC	GGTGTTC
QY	511	GCTGCTGATTCGCGAGGCTCCCACTGTTC	CGCTTAC	CGCTTAC	CGCTTAC	CGCTTAC	CGCTTAC	CGCTTAC	CGCTTAC
DB	361	GCTGCTGATTCGCGAGGCTCCCACTGTTC	CGCTTAC	CGCTTAC	CGCTTAC	CGCTTAC	CGCTTAC	CGCTTAC	CGCTTAC
QY	571	CTTCAGTTCCAAATGATGAACATCATCAAC	GGTGGCT	GGTGGCT	GGTGGCT	GGTGGCT	GGTGGCT	GGTGGCT	GGTGGCT
DB	421	CTTCAGTTCCAAATGATGAACATCATCAAC	GGTGGCT	GGTGGCT	GGTGGCT	GGTGGCT	GGTGGCT	GGTGGCT	GGTGGCT
QY	631	GTTCAGGAATTCATGCTGCTCCAAATCGGT	GCAGAC	GCAGAC	GCAGAC	GCAGAC	GCAGAC	GCAGAC	GCAGAC
DB	481	GTTCAGGAATTCATGCTGCTCCAAATCGGT	GCAGAC	GCAGAC	GCAGAC	GCAGAC	GCAGAC	GCAGAC	GCAGAC
QY	691	GGCGGGAGGCTTACCAGCACTGAAGTCCG	TCAATCA	AGGAAAGG	AGGAAAGG	AGGAAAGG	AGGAAAGG	AGGAAAGG	AGGAAAGG
DB	541	GGCGGGAGGCTTACCAGCACTGAAGTCCG	TCAATCA	AGGAAAGG	AGGAAAGG	AGGAAAGG	AGGAAAGG	AGGAAAGG	AGGAAAGG
QY	751	CTTGCGGATGAGGCGGCTTCGCTCCTTCC	GTCCGCT	GTCCGCT	GTCCGCT	GTCCGCT	GTCCGCT	GTCCGCT	GTCCGCT
DB	601	CTTGCGGATGAGGCGGCTTCGCTCCTTCC	GTCCGCT	GTCCGCT	GTCCGCT	GTCCGCT	GTCCGCT	GTCCGCT	GTCCGCT
QY	811	ATCGTTGAGGCAATCGAGAAGCTGGCTTC	ACCCGAG	CCGAGGAC	CCGAGGAC	CCGAGGAC	CCGAGGAC	CCGAGGAC	CCGAGGAC
DB	661	ATCGTTGAGGCAATCGAGAAGCTGGCTTC	ACCCGAG	CCGAGGAC	CCGAGGAC	CCGAGGAC	CCGAGGAC	CCGAGGAC	CCGAGGAC
QY	871	GACGTTGCTTCTCTGAGTTCTTCAAGGAC	GGGCTT	ACCACTT	CGAAGT	GGGCTT	ACCACTT	CGAAGT	GGGCTT
DB	721	GACGTTGCTTCTCTGAGTTCTTCAAGGAC	GGGCTT	ACCACTT	CGAAGT	GGGCTT	ACCACTT	CGAAGT	GGGCTT
QY	931	TCCGCGATGAGATGGGAAAAGTTTACGCT	GAGCTCG	TTGAGCG	GTACCC	GTACCC	GTACCC	GTACCC	GTACCC
DB	781	TCCGCGATGAGATGGGAAAAGTTTACGCT	GAGCTCG	TTGAGCG	GTACCC	GTACCC	GTACCC	GTACCC	GTACCC
QY	991	ATCGAGGACCCACTCAGGAGATGACTGG	AGGCTT	ACCACTT	CGAAGT	GGGCTT	ACCACTT	CGAAGT	GGGCTT
DB	841	ATCGAGGACCCACTCAGGAGATGACTGG	AGGCTT	ACCACTT	CGAAGT	GGGCTT	ACCACTT	CGAAGT	GGGCTT
QY	1051	GGCGACAAGGTTTCAGATCGTTGGCGAC	GACTTCT	TCGTCA	CAAC	CCCTG	AGCGCT	GAAG	1110
DB	901	GGCGACAAGGTTTCAGATCGTTGGCGAC	GACTTCT	TCGTCA	CAAC	CCCTG	AGCGCT	GAAG	960
QY	1111	GAGGCGATCGTTAAGAGCTGCCAACTCC	ATCCTC	CTGTT	TAAGT	GTAAC	CCAGAT	CGGTACC	1170
DB	961	GAGGCGATCGTTAAGAGCTGCCAACTCC	ATCCTC	CTGTT	TAAGT	GTAAC	CCAGAT	CGGTACC	1020

RESULT 9

ACC45313
ID ACC45313 standard; DNA; 1299 BP.

XX ACC45313;

AC ACC45313;

XX 17-JUN-2003 (first entry)

DT 17-JUN-2003 (first entry)

XX Modified eno DNA fragment S330E SEQ ID NO:4.

DE Fine chemical; Coryneform bacteria; Escherichia coli; microorganism;

XX genetically modified microorganism; metabolite; biosynthesis; amino acid;

KW vitamin; nucleoside; nucleotide; pigment; protein; human medicine;

KW pharmaceutical; food; animal feeding; eno; gene; ds.

XX Corynebacterium glutamicum.

OS Synthetic.

XX WO2003023016-A2.

PN 20-MAR-2003.

XX 11-SEP-2002; 2002WO-EP010174.

XX 13-SEP-2001; 2001DE-01045043.

PR (DEGS) DEGUSSA AG.

XX Farwick M, Hermann T;

XX WPI; 2003-354534/33.

XX Microorganism useful for producing e.g. fine chemicals, has permanently

PT altered phosphorylatability protein, such that biosynthesis of fine

PT chemical synthesized by microorganism is increased compared to wild-type.

XX Claim 17; Page 82; 120pp; English.

PS The present invention describes a microorganism (I), in which the

XX phosphorylatability of at least one protein has been permanently altered

CC such that the biosynthesis of at least one fine chemical synthesised by

CC the microorganism is increased compared to the wild type. Also described:

CC (1) use of a DNA (II) sequence coding for a protein which contains a

CC phosphorylation site, where the sequence contains such a mutation that

CC the protein is changed in its phosphorylatability for the production of

CC (I), or for the production of fine chemicals; and (2) a method for

CC producing fine chemicals or metabolites comprising using (I). (I) is

CC useful for producing fine chemicals or metabolites, such as amino acids,

CC vitamins, nucleosides, nucleotides, pigments or proteins. The amino acids

CC and vitamins produced using (I) can be used in human medicine, in the

CC pharmaceutical industry, food industry and in animal feeding. (I)

CC produces larger amount of desired fine chemical or a metabolite than the

CC

CC wild type. The present sequence represents a modified eno fragment DNA
 CC sequence from Corynebacterium glutamicum, which is used in an example
 CC from the present invention
 XX
 SQ Sequence 1299 BP; 275 A; 407 C; 355 G; 262 T; 0 U; 0 Other;
 Query Match 80.5%; Score 1270.2; DB 8; Length 1299;
 Best Local Similarity 99.8%; Pred. No. 5.8e-310;
 Matches 1272; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 154 GCTCAATATGACGATTCCTCGCGAATTTCTCGACTCCCGGGTAACCCACCGTC 213
 DB 10 GCTCAATATGACGATTCCTCGCGAATTTCTCGACTCCCGGGTAACCCACCGTC 69
 QY 214 GAGCAGAGGTTTTCCTGGATGACGGTTCCACCGTGTCCAGTGTTCATCCGGGCA 273
 DB 70 GAGCAGAGGTTTTCCTGGATGACGGTTCCACCGTGTCCAGTGTTCATCCGGGCA 129
 QY 274 TCCACCGGCTCCACGAGGCTCATGAGCTGCGTGAACCGTGTTCATCCGGGCAAG 333
 DB 130 TCCACCGGCTCCACGAGGCTCATGAGCTGCGTGAACCGTGTTCATCCGGGCAAG 189
 QY 334 GCGCTTTTGAAGGAGTGTGAAGAGTCAACGAGAAATCCGCGACGAGCTCGCTGGCTA 393
 DB 190 GCGCTTTTGAAGGAGTGTGAAGAGTCAACGAGAAATCCGCGACGAGCTCGCTGGCTA 249
 QY 394 GAGCTGACGATCAGCGCTCATCGACGAGCAATGATCAAGCTTGTGGCACCAGCAAC 453
 DB 250 GAGCTGACGATCAGCGCTCATCGACGAGCAATGATCAAGCTTGTGGCACCAGCAAC 309
 QY 454 AAGTCCCGCTGGGTGCAACGCAATCCTTGGTGTTCATGGCTGTTCGAAAGGCTGCT 513
 DB 310 AAGTCCCGCTGGGTGCAACGCAATCCTTGGTGTTCATGGCTGTTCGAAAGGCTGCT 369
 QY 514 GCTGATTCGCGAGGCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 573
 DB 370 GCTGATTCGCGAGGCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 429
 QY 574 CCAGTTCATGATGAACATCATCAACGGTGGCGCTCAGCTGCTGCTGCTGCTGCT 633
 DB 430 CCAGTTCATGATGAACATCATCAACGGTGGCGCTCAGCTGCTGCTGCTGCTGCT 489
 QY 634 CAGGAATTCATGATCGCTCCAATCGGTGCGAGACCTTCTCTGAGGCTCTCCGCAACGGC 693
 DB 490 CAGGAATTCATGATCGCTCCAATCGGTGCGAGACCTTCTCTGAGGCTCTCCGCAACGGC 549
 QY 694 GCGAGGCTTACACGACATGAAGTCCGTGATCAAGGAAAGGCGCTGTCCACCGACTT 753
 DB 550 GCGAGGCTTACACGACATGAAGTCCGTGATCAAGGAAAGGCGCTGTCCACCGACTT 609
 QY 754 GCGATGAGGCGGCTTCCGCTCCTTCCGCTGCTGCTGCTGCTGCTGCTGCTGCT 813
 DB 610 GCGATGAGGCGGCTTCCGCTCCTTCCGCTGCTGCTGCTGCTGCTGCTGCTGCT 669
 QY 814 GTTGAGGCAATTCGAGAGGCTGGCTTCAACCCGAGGAGGAGATCGCTTCTGCTCTGGAC 873
 DB 670 GTTGAGGCAATTCGAGAGGCTGGCTTCAACCCGAGGAGGAGATCGCTTCTGCTCTGGAC 729
 QY 874 GTTGCTTCTCTGATTTCTTCAAGGAGGCGACCTTACCACTTCCGAGGAGGAGGAG 933
 DB 730 GTTGCTTCTCTGATTTCTTCAAGGAGGCGACCTTACCACTTCCGAGGAGGAGGAG 789
 QY 934 GCAGCTGAGTGGCAAGGTTTACGCTGAGCTCGTTGACCGGTACCCCAATCGTCTCCATC 993
 DB 790 GCAGCTGAGTGGCAAGGTTTACGCTGAGCTCGTTGACCGGTACCCCAATCGTCTCCATC 849
 QY 994 GAGGACCCACTGCGAGGAGGATGATCGGAGGTTTACACCACTTCCGAGGAGGAGGAG 1053
 DB 850 GAGGACCCACTGCGAGGAGGATGATCGGAGGTTTACACCACTTCCGAGGAGGAGGAG 909
 QY 1054 GACAAGGTTTCAAGTGTGCGGAGGATTTCTTCTGCTACCAACCTTGAGCGGCTGAAGGAG 1113
 DB 910 GACAAGGTTTCAAGTGTGCGGAGGATTTCTTCTGCTACCAACCTTGAGCGGCTGAAGGAG 969

QY 1114 GGCATCGCTAAGAGGCTGCCAACTCCATCTCTGTTAAGGTGAACACGAGATCGGTACCCCTC 1173
 DB 970 GGCATCGCTAAGAGGCTGCCAACTCCATCTCTGTTAAGGTGAACACGAGATCGGTACCCCTC 1029
 QY 1174 ACCGAGACCTTTCGACGCTGCGACATGCGCTCACCGCGAGGCTACACCTCCATGATGCC 1233
 DB 1030 ACCGAGACCTTTCGACGCTGCGACATGCGCTCACCGCGAGGCTACACCTCCATGATGCC 1089
 QY 1234 CACCGTTCCGCTGAGAGCGGAGGACACCACTTCTGACCTCGAGTTGCACTCAACTGT 1293
 DB 1090 CACCGTTCCGCTGAGAGCGGAGGACACCACTTCTGACCTCGAGTTGCACTCAACTGT 1149
 QY 1294 GGCAGATCAAGACTGGTGTCTCCAGCAGCTTCCACCGTGTCCGAAAGTACAACAGCTT 1353
 DB 1150 GGCAGATCAAGACTGGTGTCTCCAGCAGCTTCCACCGTGTCCGAAAGTACAACAGCTT 1209
 QY 1354 CTCCGATCGAGCAGCTGCTTGGCGAGCGCGGCTCTACGAGGTCGCGAGGCAATTCCTCA 1413
 DB 1210 CTCCGATCGAGCAGCTGCTTGGCGAGCGCGGCTCTACGAGGTCGCGAGGCAATTCCTCA 1269
 QY 1414 CGCTTTCCAGGCTTAA 1428
 DB 1270 CGCTTTCCAGGCTTAA 1284
 RESULT 10
 ACA29854
 ID ACA29854 standard; DNA; 1275 BP.
 XX
 AC ACA29854;
 AC XX
 DT 19-JUN-2003 (first entry)
 XX
 DE Prokaryotic essential gene #11511.
 XX
 KW Antisense; ds; prokaryotic essential gene; cell proliferation;
 KW drug design; gene.
 XX
 OS Corynebacterium diphtheriae.
 XX
 FN WO200277183-A2.
 XX
 PD 03-OCT-2002.
 XX
 PF 21-MAR-2002; 2002WO-US009107.
 XX
 PR 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-032699P.
 XX
 PA (ELIT-) ELITRA PHARM INC.
 XX
 PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlson KL, Zyskind JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX
 DR WPI; 2003-029926/02.
 DR P-PSDB; ABU25984.
 XX
 PT New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX
 PS Claim 14; SEQ ID NO 17724; 1766pp; English.
 CC
 CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated

CC polypeptide or its fragment whose expression is inhibited by the
CC antisease nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisease nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
CC prokaryotic essential genes. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 1275 BP; 268 A; 378 G; 354 G; 275 T; 0 U; 0 Other;

Query Match 57.6%; Score 908.8; DB 8; Length 1275;
Best Local Similarity 82.0%; Pred. No. 9.2e-219;
Matches 1046; Conservative 0; Mismatches 229; Indels 0; Gaps 0;

QY 151 GTGGCTGAATCATGACGATATTCGCTCGGAAATTCCTGACCTCCCGGGTAACCCAAAC 210
DB 1 GTGGCTGACATATGACGATATTCGCTCGGAAATTCCTGACCTCCCGGGTAACCCAAAC 60
QY 211 GTGAGGACAGAGTTTCTCGATGACGGTTCCACCGGTGCGGAGGTGTTCCATCCGGC 270
DB 61 GTTGAGCGGAGTTTCTTGATGACGGATCCACCGGTTGCGGGGTTCTTCGGT 120
QY 271 GCATCCACCGGCTCCACGAGGCTCATGAGTGGTCAACGAGGATGCGGAGGCTGCGTGGC 330
DB 121 GCGTCCACCGGTTTACGAGGCTCAGAGCTTCGCGACGGTGGCGAGGCTACCTCGG 180
QY 331 AAGGCGTTTGAAGGAGTGAAGAGTCAACGAGGATGCGGAGGATGCGGAGGCTGCGTGGC 390
DB 181 AAGGCGTTTCAATGAGTTAAACGTTAAGGAGGATGCGTGAACGATGCGTGGT 240
QY 391 CTAGAGGCTGACGATCAGCGCTCATCGAGGAAATGATCAAGCTTGTGCGACCGCC 450
DB 241 GCAGAGCGGAGATCAGCGCTGATGATCAGCGGATGATGCTTCGAGCGGACATGAG 300
QY 451 AACAAAGTCCCGCTGGGTGCAACGCAATCTTGGTGTTCATGGCTGTTGCAAGGCT 510
DB 301 AACAAAGTCTGCTCGCGGCAACGCTATCTTGGGTGTATCATCGCGGTAGTAAAGCT 360
QY 511 GCTGCTGATCCGAGGCTCCCACTGTTCCGCTACATCGGTGACCAAGGACACAGTT 570
DB 361 GCGCAGAGTCTGCTGGCTACCTTTGACCGCTACATCGCGGCGCCCTAACGCTCAAGTT 420
QY 571 CTTCCAGTTCAATGATGAACATCATCAACGGTGGCGCTCAGCTGACTCCGGTGTGAC 630
DB 421 CTTCCAGTTCTATGATGAACATGTTAAAGGTTGGCGCACAGCTGACTCCGGGTTGAT 480
QY 631 GTTCAGAAATTCATGATGCTTCCATCGGTGACAGACCTTCTCTGAGGCTCTCCGCAAC 690
DB 481 GTTCAGAGTTCATGATGCTTCCATCGGTGCGGAGTCTTCTCTGAGGCTCTCGCATG 540
QY 691 GCGCGGAGGTTTACGAGGCTCAAGTCCGTCATCAAGGAAAGGGGCTGTCACCGGA 750
DB 541 GGTGAGAGGTTTACCACTCTTGAAGTCCGTTGATTAAGTCCAAAGGGGACTTCCACCGG 600
QY 751 CTTGGCGATGAGGCGGCTTCGCTCTTCCGCTCGGCTCCACCGCTGAGGCTCTTGACCTT 810

DB 601 CTGCGGACGAAGGTGGTTTCGCACCTTCTGTTGAGTCCACCAAGGAGCTCTCGACCTC 660
QY 811 ATCGTTGAGCAATCGAGAAAGGTGGCTTACCCCGAGGAAAGGACATCGTCTTGTCTGTG 870
DB 661 ATCGTTGAGCAATCGAGAAAGGTGGCTTCAAGCCAGGTGCTGACATCGCACTTGCACCTC 720
QY 871 GAGTTGCTTCCCTGAGTTCTTCAAGGACGGCACCTACCACTTCAAGGTTGGCCAGCAC 930
DB 721 GAGTTGCTTCCCTGAGTTCTTCAAGGATGGCAAGTACCACTTCAAGGTTGGCCAGCAC 780
QY 931 TCCGCAAGCTGAGATGGCAAGCTTTCGCTGAGTCTGTTGACCGGTACCCCAATCGTCTCC 990
DB 781 ACCGCTGAGGAGATGGCAAGCTTACGAGCAGTCTCATCGTGAATACCAATTTGTTCC 840
QY 991 ATCGAGACCCATCGCAGGAAGATGATGGGAGGTTTACCAACCTTACCGCAACCATC 1050
DB 841 ATCGAGACCCATCGCAGGAAGACGACTGGGAGGGCTTACACCGGCTTGAACCGGCAATC 900
QY 1051 GCGCAAGAGTTGAGTTCGTTGGGAGGACTTCTTGTGTCACCAACCTGAGGCGCTGAAG 1110
DB 901 GGTGACAAGTTTCAAGTTCGTCGCGGAGCTTCTTGTGTCACCAACCGGAGCTCAAG 960
QY 1111 GAGGCGATCGCTAAGAGGCTGCCAATCTCCATCTGTTAAGGTGAACCAAGATCGGTACC 1170
DB 961 GAAGGCAATCGAAGAGGCTGCCAATGCTTGTGTTAAGGTGAACCAAGATCGGTACC 1020
QY 1171 CTACCGAGACCTTCGACGCTGTCGATCGGCTACCGCGGAGGCTACACTCCATGATG 1230
DB 1021 CTGACCGAGACCTTCGACGCTGTTGATCTCGCACACCGCAACCGGCTACCGGACCATG 1080
QY 1231 TCCACCGTTCCGTTGAGCGGAGGACACCACTTCTGCTGACCTCGCAGTTGCACTCAAC 1290
DB 1081 TCCACCGTTCCGCGGAGACTGAAGACACCACTTGTGCTGACCTTGCAGTTCGCAATTGGGC 1140
QY 1291 TGTGGCCAGATCAAGACTGCTGCTGTCAGACAGCTTCCGACCGTGTGCAAAAGTCAACAC 1350
DB 1141 TGTGGCCAGATCAAGACTGCTGCTGTCAGACAGCTTCCGACCGGCTTGCAGGACATCAG 1200
QY 1351 CTTCTCCGATCGAGCAGCTGCTTGGCGAGCGCGGCTCTACGAGGTGCGAGCGCATTC 1410
DB 1201 CTTCTGCGCATCGAGCAGCAGCTTGAAGTGCAGAGCTTACGCGAGCGGCTTACGACATTT 1260
QY 1411 CCAGCTTTTCAGGCG 1425
DB 1261 CCAGCTTTTCAGGCG 1275

RESULT 11
ADS56353
ID ADS56353 standard; cDNA; 1269 BP.
XX
AC ADS56353;
XX
DT 02-DEC-2004 (first entry)
XX
DE Bacterial polynucleotide #8340.
XX
KW Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polynucleotide; gene; ss.
XX
OS Bacteria.
XX
PN US2003233675-A1.
XX
PD 18-DEC-2003.
XX
PF 20-FEB-2003; 2003US-00369493.

XX 21-FEB-2002; 2002US-0360039P.
 PR (CAOY/) CAO Y.
 PA (HINK/) HINKLE G J.
 PA (SLAT/) SLATER S C.
 PA (CHEN/) CHEN X.
 PA (GOLD/) GOLDMAN B S.
 XX
 PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
 XX WPI; 2004-061375/06.
 DR
 XX
 XX New recombinant DNA construct comprising a promoter positioned to provide
 PT for expression of a polynucleotide encoding a polypeptide from a
 PT microbial source, useful for producing plants with improved properties.
 XX
 XX Claim 1; SEQ ID NO 32027; 122pp; English.
 XX
 CC The invention relates to a recombinant DNA construct comprising a
 CC promoter functional in a plant cell, where the promoter is positioned to
 CC provide for expression of a polynucleotide encoding a polypeptide from a
 CC microbial source. The invention also relates to a transformed plant
 CC comprising the recombinant DNA construct and a method of producing a
 CC transformed plant having an improved property. The plant is a crop plant
 CC such as maize or soybean. The method of producing a transformed plant
 CC having an improved property comprises transforming a plant with the
 CC recombinant DNA construct and growing the transformed plant, where the
 CC polynucleotide or polypeptide is useful for improving plant properties.
 CC The recombinant DNA construct is useful for producing plants with
 CC improved plant properties, e.g. improved cold, heat or drought tolerance,
 CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
 CC increased resistance to plant disease, better growth rate by modification
 CC of the cell cycle pathway with plant growth regulators, increased rate of
 CC homologous recombination, modified seed oil or protein yield and/or
 CC content, improved yield by modification of carbohydrate, nitrogen or
 CC phosphorus use and/or uptake, by modification of photosynthesis or by
 CC providing improved plant growth and development under at least one stress
 CC condition, improved lignin production or improved galactomannan
 CC production. This sequence represents a bacterial polynucleotide used in
 CC the scope of the invention. Note: The sequence data for this patent did
 CC not form part of the printed specification but was obtained in electronic
 CC format from USPTO at seqdata.uspto.gov/sequence.html.
 XX
 SQ Sequence 1269 BP; 227 A; 421 C; 443 G; 178 T; 0 U; 0 Other;

QY 512 CTGCTGATTCGCGAGGCTCCACCTGTTCCGCTACATCGGTGGACCAACACGACGTTTC 571
 Db 362 CCGCGAGAGCGCGACCTGCGGCTCTTCCGCTCTTCCGCTGCGGCGCCGACGCGACGTCG 421
 QY 572 TTCCAGTTTCCAAATGATGAACATCATCAACGGTGGCGCTCAGCGTGACTCCGGTGTGACG 631
 Db 422 TGGCGGTGCCGATGATGAACATCTCAACGGGGGTGGCGACGCGACACCAACGTCGACA 481
 QY 632 TTCAGGAATTCATGATCGCTCCATCGGTGCGAGACCTTCTCTGAGGCTCTCCGCAACG 691
 Db 482 TCCAGGAGTTTCATGATCGCCCCCATCGGGGCGGAGAGCTTCCGGAGGCGCTCGCCTGGG 541
 QY 692 GCGCGAGGCTTACACGACCTGAAGTCCGTCATCAAGGAAAGGCGCTGTCCACCGGAC 751
 Db 542 GCGCGAGGCTTACACCTCGCTCAAGGGGCTGCTCAAGCCCAAGGCTTGGCCACCGGTG 601
 QY 752 TTGCGATGAGGGCGGCTTCCGCTCTTCCGCTGCGCTCCACCCGTGAGGCTCTTGACCTTA 811
 Db 602 TCGGTGACGAGGCGGGTTTCGCGCCCAACCTGGACAGCAACCGCGCGCTCGACCTGA 661
 QY 812 TCGTTGAGGCAATCGAAGGCTGGCTTCAACCCAGGCAAGGACATCGCTCTTGTCTGCG 871
 Db 662 TCAGCGAAGCCATCACGAAGCGCGGCTTCAAGCTGGGCGGAGACATCGCTTGGCTTGG 721
 QY 872 ACCTTGTCTCTCTGAGTTCTTCAAGGACGCGACCTACCACTTCGAAGGTGGCCAGCACT 931
 Db 722 ACCTTGTCTCTCTGAGTTCTTCAAGGACGCGACCTACCACTTCGAAGGTGGCCAGCACT 781
 QY 932 CCGCAGCTGAGATGGCAAAAGCTTTACGCTGAGCTCGTTGACGCGTACCAATCTGCTCCA 991
 Db 782 GCGCGAGGAGATGCGCGCTTACTACCGAGCTGGTTCGAGTCTTACCGCTGGTGTGCA 841
 QY 992 TCGAGGACCACTCGAAGGATGATGCTGGGAGGTTTACCACTTACCGCAACCACTACCG 1051
 Db 842 TCGAGGACCACTCGAAGGATGATGCTGGGAGGTTTACCACTTACCGCAACCACTACCG 901
 QY 1052 GCGACAAGTTTACAGATCGTTGGCGGACGACTTCTTCTGCTACCAACCTGAGCGCTGAAG 1111
 Db 902 GCGACCGGCTGCAACTGCTGGCGGACGACTTCTTCTGCTACCAACCTGAGCGCTGACG 961
 QY 1112 AGGCATCGCTTAAGAAGGCTGCCAACTCCCTCTGTTAAGGTGAACAGATCGGTACCC 1171
 Db 962 GCGGTATCGAGGCGGCTGCCAACTCGCTGCTGGTCAAGTGAACAGATCGGTACGC 1021
 QY 1172 TCACCGAGACTTCGACGCTGTGCAATGGCTTACCGCGCAGGCTTACCACTTCCATGATGT 1231
 Db 1022 TCACCGAGACTTCGACGCTGTGCAATGGCTTACCGCGCAGGCTTACCACTTCCATGATGT 1081
 QY 1232 CCACCGCTTCGCTGAGACCGGACACCACTTCTGCTGACCTCGCAGTTTGCATCACT 1291
 Db 1082 GCGACCGGCTTCGCTGAGACCGGACACCACTTCTGCTGACCTCGCAGTTTGCATCACT 1141
 QY 1292 GTGCCAGATCAAGACTTGGTGTCTCCAGACGCTTCCGACCGTGTGCGAAAGTACAAACG 1351
 Db 1142 CGGGCAGATCAAGACGCTGCTCCCGCGCGCAGCAGCGGCTCGCCAGTACAAACGAG 1201
 QY 1352 TTCTCCGATCGAGCAGCTGTTGGCGACGCGCGGCTTACGCGAGGTCGAGCGCATTTCC 1411
 Db 1202 TGCTCGGATCGAAGAAGAGCTCGACGCGAGCGCTCTACGCGGAGGCAACGCGTTCC 1261
 QY 1412 CACGCTT 1418
 Db 1262 CGGTTT 1268

RESULT 12
 ACA38079
 ID ACA38079 standard; DNA; 1287 BP.
 XX
 XX ACA38079;
 XX
 DT 19-JUN-2003 (first entry)

XX Prokaryotic essential gene #19735.
 XX
 DE Antisense; ds; prokaryotic essential gene; cell proliferation;
 XX drug design; gene.
 XX
 XX Mycobacterium avium.
 OS
 XX W0200277183-A2.
 PN
 XX 03-OCT-2002.
 PD
 XX 21-MAR-2002; 2002WO-US009107.
 PF
 XX 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 PR
 XX (ELIT-) ELITRA PHARM INC.
 PA
 XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 PI
 XX WPI; 2003-029926/02.
 DR P-PSDB; ABU34209.
 DR
 XX New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 PT
 XX Claim 14; SEQ ID NO 25949; 1766pp; English.
 PS
 XX The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid;
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
 CC prokaryotic essential genes. Note: the sequence data for this patent did
 CC not form part of the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 XX Sequence 1287 BP; 214 A; 424 C; 475 G; 174 T; 0 U; 0 Other;
 SQ
 Query Match 41.0%; Score 647; DB 8; Length 1287;
 Best Local Similarity 69.9%; Pred. No. 8.8e-153;
 Matches 888; Conservative 0; Mismatches 380; Indels 3; Gaps 1;
 151 GTGGCTGAATCATGCACGTATTCGTCGCGAAATTCGACTCCCGCGGTAACCCCAACC 210

1 GTGCCGATATTCGAGCAGGTCGGGGCCCGGAGATCCTCGACTCCCGCGGTAAACCCGACA 60
 211 GTGAGGCGAGAGGTTTCTTGATGAGCGTTCACGAGTGTCCAGGTGTCCAGGTGTTCATCCGGC 270
 61 GTGAGGTGAGATCGCTTCGCGGAAACGTTCCCGCGCGCGGTGCGGTCCGG 120
 271 GCATCCACCGGCGTCCACGAGGCTCATGAGTCGCTGAGCGTGGCGGATCGCTACCTACCTGGGC 330
 121 GCCTGACCGGTGAGCAGGAGGCTGCGAACTCGCGACCGCGGGAGCGGTACGGCGGC 180
 331 AAGGCGGTTTGAAGGCGAGTTGAAAACGTCACAAAGAAATCGCGCACGAGCTCGCTGGC 390
 181 AAGGCGGTGAGAGGCGGTGAGGCGGTGCGACGAGATCGCGCGCGGTGATCGG 240
 391 CTAGAGGCTGAGATCAGCGCCCTCATCGACGAAGCAATGATCAAGCTTGATGGACCGCC 450
 241 CTGAACCGCGACACGAGCGCTGCTGGTGGACAGCGCTGCTGACCTTGGACCGCACCCCG 300
 451 AACAACTCCGCTGGGTGCAACGCAATCCTTTGGTGTTCATGGCTGTTGCAAAAGGCT 510
 301 GACAAGTCGAGGCTGGGCGCAACGCGATCCTGGGCGTGTGCTGGCGGTGGCGCAAGGCG 360
 511 GCTGCTGATTCGCGAGCGCTCCCACTGTTCCGCTACATCGGTGAGCAACGACACAGTT 570
 361 GCCGCGATTCGCGGAGTTGCGCTGTTCCGCTACCTGGCGGGGCCCAACGCGCACATC 420
 571 CTTCCAGTTCATGATGAAATCATCAACGTTGGGCTCACGCTGACTCCGCTGTGTGAC 630
 421 CTGCGGTGCGATGATGAACATCCTCAACGCGGTGCGCACACCGCGCTCGAC 480
 631 GTTCAGGAATTCATGATCGCTCCAATCGGTGACAGACCTTCTCTGAGGCTCTCCGCAAC 690
 481 ATCAGAGATTCATGTCGCGCGCGAGTTTCGCGAGGATTCGCGGATTCGCGTGG 540
 691 GCGCGAGGCTTACCAACGACATGAAAGTTCGTCATCAAGGAAAGGCGCTGTCCACCGA 750
 541 GGTGCGAGGTGATCACTCGTGAAGTCGCTGTAAGAAAGAGGAGGCGCTGAGCACCGC 600
 751 CTTGGCGATGAGGCGGCTTCGCTCCTTCGTCGGCTCCACCGTGAGGCTCTTGACCTT 810
 601 CTGCGGACAGAGGCGGCTTCGCGCGGAGCTGCGCGGCGCACACCGCGGCGTGAGCCTG 660
 811 ATCGTTGAGCAATCGAGAGGCTGGCTTCAACCGGCAAGGAGACATCGCTTCTGTCTG 870
 661 ATCGCGCGGCGCATCGAATCGGCGGCTTCAAACTCGGACCGAGCTGGCGCTGGCCCTC 720
 871 GAGCTTGCTTCTCTGAGTTCTTCAAGGACGGCAC---CTACCACTTCGAAGGTGGCCAG 927
 721 GACGCGCGGCCACCGAGTTCTACAGCGACGGCACCGGCTACAAGTTTCGAGGGGAGCACC 780
 928 CACTCCGACGTGAGATGGCAACAGTTTACGCTGAGCTCGTTGACGCTACCCCAATCGTC 987
 781 CGCACCGCGGAGAGATGGCGGAGTTTACGCGGGGCTGCTCGCGCGGTATTCGCTGGTG 840
 988 TCAATCAGGACACCGCTGCGAGGAGATGACTGGAGGAGTTTACACCAACCTCACCGCAACC 1047
 841 TCAATCGAAGACCGCTGTCGAGATGATGGACGCGCTGGGCGGCGCTGACCGCGTGG 900
 1048 ATCGCGCAAGGTTTCAAGTGTGGCGAGCATTTCTTTCGTCACCAACCTTGAGCGGCTG 1107
 901 ATCGCGACCGGGTGCAGCTCGTGGCGAGCATCTTTCGTCAGCAACCCCGCAACGCGCTG 960
 1108 AAGGAGGCGATCGCTAAGAGGCTGCCAATCCATCTGTTAGGTGAACCAAGATCGGT 1167
 961 GAAGAGGTTATCGGAGAGGCGCTCGAAATTCGTTGCTGTTCAAGGTGAATCAAGTCGGC 1020
 1168 ACCCTCACCGAGACTTTCGACGCTGTGAGATGGCTCACCGCGAGGCTACACCTCCATG 1227
 1021 ACGTGTACCGAGAGCGCTGGAGCGCGTTCGCGTGGCCCGCACACAGCGGCTACCGCACATG 1080
 1228 ATGTCCACCGTTCGCGTGAGACCGGAGACACCACTTGTGACCTGCGAGTTGCACTC 1287

Db 1081 ATGAGCCACCGCAGCGGTAAACACGAGACACACGATCGCCGACCTGGCGGTGGCCGTC 1140
 Qy 1288 AACTGTGGCGAGATCAAGACTGTGCTCCAGCAGCTTCGACCGCTGTGCGAAAGTACAAC 1347
 Db 1141 GGACGCGGAGATCAAGACCGGAGCGCGCGCGCGAGCGGGTGGCCNAAGTACAAC 1200
 Qy 1348 CAGTTCTCCGCATCGAGCAGCTCTGTTGGCGACCGCGCGGTCTACGAGGTCGCGAGCGCA 1407
 Db 1201 CAGCTGTGCGGATCGAGGAGCGCTCGCGCAGCGCGCGGTACGCGCGCGACCTGGCC 1260
 Qy 1408 TTCCACGCTT 1418
 Db 1261 TTCCCGCGGT 1271

RESULT 13

AA559564/C
 ID AA559564 standard; DNA; 30159 BP.

AC AA559564;

XX 13-FEB-2002 (first entry)

XX Propionibacterium acnes immunogenic protein encoding DNA #59.

DE SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
 KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 KW dermatological; osteopathic; neuroprotectant; ds.

XX Propionibacterium acnes.

OS WO200181581-A2.

XX 01-NOV-2001.

XX 20-APR-2001; 2001WO-US012865.

XX 21-APR-2000; 2000US-0199047P.

XX 02-JUN-2000; 2000US-0208841P.

XX 07-JUL-2000; 2000US-0216747P.

XX (CORI-) CORIXA CORP.

XX Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;

XX L'maisonneuve J, Zhang Y, Jen S, Carter D;

XX WPI; 2001-616774/71.

XX Propionibacterium acnes polypeptides and nucleic acids useful for
 PT vaccinating against and diagnosing infections, especially useful for
 PT treating acne vulgaris.

XX Claim 1; SEQ ID NO 59; 1069pp; English.

XX Sequences AA559506-AA559804 represent DNA molecules encoding
 CC Propionibacterium acnes immunogenic polypeptides. The proteins and their
 CC associated DNA sequences are used in the treatment, prevention and
 CC diagnosis of medical conditions caused by P. acnes. The disorders include
 CC SAPHO syndrome (synovitis, acne, pustulosis, hypertosis and
 CC osteomyelitis), uveitis and endophthalmitis. P. acnes is also involved in
 CC infections of bone, joints and the central nervous system, however it is
 CC particularly involved in the inflammatory lesions associated with acne
 CC vulgaris. A method for detecting the presence or absence of P. acnes in a
 CC patient comprises contacting a sample with a binding agent that binds to
 CC the proteins of the invention and determining the amount of bound protein
 CC in the sample. The polypeptides may be used as antigens in the production
 CC of antibodies specific for P. acnes proteins. These antibodies can be
 CC used to downregulate expression and activity of P. acnes polypeptides and
 CC therefore treat P. acnes infections. The antibodies may also be used as
 CC diagnostic agents for determining P. acnes presence, for example, by
 CC enzyme linked immunosorbent assay (ELISA). This sequence encodes the
 CC polypeptides shown in AAU54183-AAU54498 and AAU67555-AAU67557. Note: The

CC sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 30159 BP; 6067 A; 9304 C; 8907 G; 5878 T; 0 U; 3 Other;
 Query Match 40.7%; Score 642.4; DB 4; Length 30159;
 Best Local Similarity 69.3%; Pred. No. 3.6e-151;
 Matches 890; Conservative 0; Mismatches 391; Indels 3; Gaps 1;
 Qy 152 TGGCTGAAATCATGCACGTATTCCGCTCGCAAAATCTCGACTCCGGGGTAACCCACCG 211
 Db 1539 TGGCAACCATCGAATTCATCGAAGCCGTCGAGATTCCTCGATTCCTCGGGCAACCCGACCG 1480
 Qy 212 TCGAGGCAGAGGTTTTCTCTGATGACGGTTCCCAACGGTGTGCGAGGTGTTCCATCCGGCG 271
 Db 1479 TTGAGGTCGAGATGATCTCGACGACGCGACCCAGGCCGGCTGCGGTTCTTCGGGGCG 1420
 Qy 272 CATCCACCGCGCTCCACGAGGCTCATGAGCTGCGTGACGG---TGGCGATCGCTACCTGG 328
 Db 1419 CTTGACCCGTCAGTTCGAGGCCGTTGAGCTGCGTGACGGCGATAAGAAGCGTTACTCCG 1360
 Qy 329 GCAAGGCGTTTTGAAGCGAGTTGAAAACGTCAACGAAGAAATCGGACGAGCTCGCTG 388
 Db 1359 GTAAGGGTGTCTCAAGGCTGTTGAGAACGTCACGAGAAGATCGCCGAGGAGTGTCTG 1300
 Qy 389 GCCTAGAGGCTGACGATCAGCGCTCATCGACGAAGCAATGATCAAGCTTCATGCGACCG 448
 Db 1299 GCTGTGACGAGTGACGCGCATTTATCGACCATGATCATGATCGAGCTCGACGATCCG 1240
 Qy 449 CCAACAAGTCCCGCTGGGTGCAAAACGAATCTTGGTGTGTTTCCATGGCTGTTCACAAAG 508
 Db 1239 ACAACAAGGCAAGTTGGGTGCTAAACGCCATCTTGGTGTCTCCCTGGCTGCTCATG 1180
 Qy 509 CTGCTGCTGATTCGCGAGGCTCCCACTGTTCCGCTACATCGGTGCGACCAACGACACG 568
 Db 1179 CTGCTGCTGATGTGCGGAGCTCCCGCTGTACCACTACCTCGGGCGAGCCGAATCCCAAG 1120
 Qy 569 TTCTTCCAGTTCCAATGATGAACATCATCAACGGTGGCGCTCAGCTGACTCCCGTGTG 628
 Db 1119 TGTGCCCCGTTCCAATGATGAACATCTCAACGGTGGTGGCCACGCGATTTCGACGTTG 1060
 Qy 629 ACGTTCAGGAATTCATGATCGCTCCAATCGGTGCGAGAGACCTTCTCTGAGGCTCTCCGCA 688
 Db 1059 ACATCCAGGATTCATGATTCGCCCGATCGGCGGGAATCCTTTAAGCAGGCTACAGAGT 1000
 Qy 689 ACGGCGGAGGCTTACACGCACTGAAGTCCGTCATCAAGGAAAGGCGCTGTCCACCG 748
 Db 999 GGGGCGCTGCGGTTTACCACTCCCTCAAGAAGGTTCTCAAGGACAAAGGGCTTGGCTACCG 940
 Qy 749 GACTTGGCGATGAGGGGGCTTCGCTCTTCCGTCGGCTCCACCGTGAGGCTCTTGACC 808
 Db 939 GTCTGGGTGACGAGGGCGGTTTCGCCCCCAACCTTCCGACGACGCGCGCTTTGGACC 880
 Qy 809 TTATCTGTGAGGCAATCGAGAAGGCTGGCTTCACCCCGAGGCAAGAGATCCTCTTGTGTC 868
 Db 879 TCATCTTGAAGCATCAAGGCCGCGGCTTCGAGCGGGTAAGGAGCTGCGCCCTCGGCC 820
 Qy 869 TGAAGTGTGCTTCTCTGAGTTCTTCAAGGACGCACTACCACTTCGAAGGTGGCCAGC 928
 Db 819 TTGATGTGCTGCTTCGAAATCTTCGAGGACGGCAAGTACACGTTTCGAGGGCCAGGCTA 760
 Qy 929 ACTCCGAGCTGAGATGGCAACGTTTACGCTGAGCTGTTGAGCGGTACCAATCGTCT 988
 Db 759 AGACCTCGGCTGAGATGATCGCGTACTACGAAGGCCCTCATGCGCAAGTATCCGCTGGTCT 700
 Qy 989 CCATCGAGGACCCACTCGAGAAGATCACTGGGAGGGTTACACCAACTCACCAGCAACCA 1048
 Db 699 CTATTGAGGATCCGTTGACGAGGAGGATTTGGGACGGTTGGGCCGAGTTACGAGAGAGC 640
 Qy 1049 TCGGCGCAAGGTTTCAAGTGGGACGACTTCTTCGTTCACCAACCTTGAGCGCCCTGA 1108
 Db 639 TTGTTGAGAAGATCCAGATCGTTGGCGACGACTTGTTCGTCCTCACTAAACCCGAGCGCTTG 580

Db 819 TTGATGTGCTGCTTCCGAATTTCTCGAGGACGGCAAGTACACGTTCTGAGGGCCAGGCTA 760
 QY 929 ACTCCGAGCTGAGATGCGAAACGTTTACGCTGAGCTGCTTACGCGTACCCAACTGCT 988
 Db 759 AGACCTGGCTGAGATGATCGGTACTACGAGGCTTATCGCCAAAGTATCCGCTGCT 700
 QY 989 CCATCGAGGCCACTCGAGGAAGATGACTGGGAGGTTACCAACCTCACCCCAACCA 1048
 Db 699 CTATTGAGGATCCGTTGGACGAGGAGATTGGGACGCTTGGCCGAGTTTACGAAGAAGC 640
 QY 1049 TCGGCGACAGGTTTCAGATCGTTGGCGACGACTTCTTCGTCACCAACCTGAGCGCTGA 1108
 Db 639 TTGGTGAAGATCCAGATCGTTGGCGACGACTTGTTCGTCACCTAACCCGGAAGCGCTTG 580
 QY 1109 AGAGGGCATCGCTAAGAAGCTGCCAACTCCATCTCTGGTTAAGGTGAACAGATCCGTA 1168
 Db 579 CCAAGGAATCGAGACCAAGGCTGCCAAGCCCTGCTCGTCNAGGTGAACAGATTGGCT 520
 QY 1169 CCCTCACCGAGACTTCGACGCTGTGACATGGCTCACCGCGAGGCTTACACCTCCATGA 1228
 Db 519 CCCTCTCGGAGACTATCGACCGCTCGAGCTGGGCCATCGCAACGGCTACCGCTGCATGA 460
 QY 1229 TGTCCACCGTTCGCTGAGACCGAGGACACCACTTGTGACCTGCGAGTTGCATCA 1288
 Db 459 TGTCCACCGTTCGCTGAGACCGAGGACCACTATCGCGGATCTCGCGCTCGCGTAT 400
 QY 1289 ACTGTGCCAGATCAAGACTGGTGCTCCAGCACGTTCCGACCGGTGCGAAAGTACAAAC 1348
 Db 399 CGACTGTGAGATTAAATCCGGTGCCCGCCCGCGGAGCGCATCGCCNAGTACAAAC 340
 QY 1349 AGCTTCTCCGATCGAGCAGCTGTGGCGACGCGCGGCTGTACGACAGGTGCGAGCGCAT 1408
 Db 339 AGTTGCTGCGTATTGAGGAGGAATGGGCGACTCGGCGGAGTACGCGGTGCTTCGCTT 280
 QY 1409 TCCACGCTTTCAGGGCTAAATAA 1432
 Db 279 TCCGCGCTTCCAGGCTGACTGA 256

RESULT 15

ABX09144
 ID ABX09144 standard; DNA; 45190 BP.

XX AC ABX09144;

XX DT 08-APR-2003 (first entry)

XX DE Mycobacterium tuberculosis H37Rv BAC clone BAC-Rv267.

XX KW Mycobacteriosis; survival; virulence; protective antigen; vaccine;
 XX KW mycobacterial disease; tuberculosis; leprosy; ds; cosmid.

XX OS Mycobacterium tuberculosis.

XX XX W0200274903-A2.

XX XX 26-SEP-2002.

XX PF 22-FEB-2002; 2002MO-IB001973.

XX PR 22-FEB-2001; 2001US-0270123P.

XX PA (INSP) INST PASTEUR.

XX PI Cole S;

XX DR WPI; 2002-759885/82.

XX PT Identifying and selecting genes for survival or virulence of mycobacteria
 XX PT by a comparative genomic analysis of the sequences of Mycobacterium
 XX PT tuberculosis and M. lepreae.

XX PS Disclosure; Fig 8; 874pp; English.

XX

CC This invention relates to a novel method for identifying essential genes
 CC for survival or virulence of mycobacteria species. The method comprises
 CC aligning the genomic sequence of a first mycobacterium species on a
 CC genomic sequence of a second mycobacterium species and selecting a
 CC polynucleotide sequence that is highly conserved in both genomes with no
 CC counterparts in other bacterial genomic sequences and that corresponds to
 CC an essential gene for the survival or virulence of mycobacterium species.
 CC The method of the invention is useful for detecting M. tuberculosis or M.
 CC lepreae infection. The method reduces the number of potential new targets
 CC and protective antigens for new drugs and vaccine compositions to treat
 CC and prevent mycobacterial diseases, particularly tuberculosis and
 CC leprosy. The present sequence represents a Mycobacterial cosmid DNA
 CC sequence used in the method of the invention

SQ Sequence 45190 BP; 7490 A; 14522 C; 15204 G; 7974 T; 0 U; 0 Other;

Query Match 40.1%; Score 632.6; DB 6; Length 45190;
 Best Local Similarity 68.4%; Pred. No. 1.2e-148;
 Matches 891; Conservative 0; Mismatches 409; Indels 3; Gaps 1;

QY 140 AGGAGGCCACAGTGGCTGAATCATGACGTAATTCGCTCGGAAATTTCTGACTCCCGCG 199
 Db 19931 AGGAGAACCCAGTGCCTGATTTATCGAGCAGGTTAGGGCCCGAGAGATCCTCGATTCCCGCG 19990
 QY 200 GTAAACCCAAACCGTCCGAGCGAGAGGTTTTCTTGGATGACGGTTCCACGGTGTCCGAGGTG 259
 Db 19991 GCAACCCGACGGTGGAGTTCGAGTGGCGCTTATTCGACGGGACATTCGCCCGGCGCGG 20050
 QY 260 TTCCATCCGCGCATCCACCGGCTCCACGAGGCTCATGAGTCTGCTGACGGTGGCGATC 319
 Db 20051 TGCGTCCGCGCTCCACCGGCGAGCAGAGGCGCTCGAGTTGCGGACGCGCGCGATC 20110
 QY 320 GCTACCTGGGCAAGGGCGTTTTGNAAGCAGTTGAAACGTCACAGAGAAATCCGCGACG 379
 Db 20111 GCTACGGCGCAAGGGCTGCAAAAGCGCTGTCAGGCTGTTCTTGTGAGATCGGCGCG 20170
 QY 380 AGCTCGCTGCTAGAGGCTGACGATCAGCGCTCATCGACGAAGCAATGATCAAGCTTG 439
 Db 20171 CCGTCATCGACTCAACCGCGACGACGACGAGTGGTTCGACGCGCTGGTGACCTAG 20230
 QY 440 ATGGCACCGCAACAGTCCCGCTGGGTGCAACGCAATCTTTGGTGTTCATGGGTG 499
 Db 20231 ACGCACCCCGCAAGTCCCGCTGGGCGCAACGCGATCTTGGGTGTCTCGCTCGCTG 20290
 QY 500 TTGCAAGGCTGCTGCTGATTCCGACGCTCCCACTGTTCCGCTACATCGCTGACCAA 559
 Db 20291 TTGCCAAGCGCGCGCGGATTCGCGGAGCTGCTTCCGTTATGTGCGGGGGCCAA 20350
 QY 560 ACGCACAGCTTCTCCAGTTTCCAAATGATGAACATCATCAACGCTGGCGCTCACGCTGACT 619
 Db 20351 ACGGACATCTTCCCGGTACCGATGATGAACATCTTCACGCGGCGGACACCGCGATA 20410
 QY 620 CCGGTGTGACGTTCAGGAATTCATGATCGCTCCAATCGGTGCGAGAGACCTTCTCTGAGG 679
 Db 20411 CCGCTGTGACATTCGAAGAGTTTCAATGGTGGCGCAATTGGCGCGCCCGAGCTTCGTCGAGG 20470
 QY 680 CTCTCCGAACGCGCGGAGGTCTACCAAGCAGTGAAGTCCGTATCAAGAGAAAGGGCC 739
 Db 20471 CGTTGCGCTGGGCGCTGAGGTGTACCAAGCTGCTCAAGTCCGTTCTGAAAAAGAGGGGC 20530
 QY 740 TGTCCACCGGACTTGGCGGATGAGGGCGCTTCCGCTCTCTCGCTCGGCTCCACCGTGGG 799
 Db 20531 TGTCCACCGGCTGGGCGACGAAGGCGCTTCCGCCCGGATGTGCGCGGACACCGCGG 20590
 QY 800 CTCTTGACCTTATCGTTGAGGCAATCGAAGGCTGGCTTTCACCCCGGCAAGACATCG 859
 Db 20591 CGTTGACCTGATCAGCGCGGCGCATCGAGTCCGCGGCTTGCAGCAGCCCGCGCGAGCTG 20650
 QY 860 CTCTTGCTTGAGAGTGTCTCTCTGAGTTCTTCAAGGACGGCAC---CTACCACTTCG 916
 Db 20651 CGTTGGCCCTGGACGCGCGCGGCCACCGAGTTCTTTCACCGGACCGGCTACGTTCTTCG 20710

Qy	917	AAGGTGCCAGCACTCCGACGCTGAGATGGCAACGTTTACGCTGAGCTCGTTGACGCGT	976
Db	20711	AGGGCACCAACCCGTACCCGACAGACGAGATGACCGAGTTCTACGGGGGCTCTCGGCGCCT	20770
Qy	977	ACCCAATCGTCTCCATCGAGGACCCACTGCAGGAAGATGACTGGGAGGGTTACACCAACC	1036
Db	20771	ACCGCTGGTGTGATCGAAGACCCACTGTCCGAGACGATTGGGACGGCTGGGCGCGC	20830
Qy	1037	TCACCGCAACATCGGCGACAAGGTTGAGTTCGTTGGCGAGACTTCTTCGTCAACCAACC	1096
Db	20831	TGACGGCTCGATCGGTGACCGGTGCAATCGTCGGCGACGACATCTTTGTCAACCAATC	20890
Qy	1097	CTGAGCGCCTGAAGGAGGCGATCGCTAAGAGGCTGCCAACTCCATCCTGTTAAGGTGA	1156
Db	20891	CCGAGCGGCTCGAGGAGGGCATCGAACGGGCGTGGCAATGCGTTGCTGTCAGGTGA	20950
Qy	1157	ACCAGATCGGTACCCCTCACCGAGACCTTCGACGCTGTCGACATGGCTCACCGGCGAGGCT	1216
Db	20951	ACCAGATCGGACGTTGACCGAGACACTCGACGCGGTACGCTGGCTCACCGGCGGAT	21010
Qy	1217	ACACCTCCATGATGCCACGTTCCGGTGAGACCGAGGACACCACTTGTGACCTCG	1276
Db	21011	ACCGACGATGATCAGTCACCGCAGTGGCGAGACGAGGACACCATGATCGCGACCTCG	21070
Qy	1277	CAGTTGCACTCAACTGTGGCCAGATCAAGACTGGTGTCTCCAGACACTTCCGACCGTCTCG	1336
Db	21071	CGGTGGCCATCGGACGCGGGCAGATCAAGACGGGCGGCTGCTCGCAGTGAGCGCGTCG	21130
Qy	1337	CAAAAGTACAACCAAGCTTCTCCGCAATCGAGCAGCTGCTTGGCGACCGCGGCTCTACGCAG	1396
Db	21131	CANAATACAACCAAGCTGCTGCGGATCGAAGAGGCGCTTGGCGACGCGCGCGCTACGCGG	21190
Qy	1397	GTCGAGCGCATTCGCCAGCTTTCCAGGCTTAAATAAAGCGCT	1439
Db	21191	GCGACCTGGCATTTCTCGGTTCCGTCGTCGAGACGAAATAGGT	21233

Search completed: September 28, 2005, 17:01:31
Job time : 801.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 28, 2005, 16:03:14 ; Search time 4756.5 Seconds
(without alignment)
12628.070 Million cell updates/sec

Title: US-10-728-947-1

Perfect score: 1578

Sequence: 1 ggctgggatatgggtagt.....ctcaagcagggaacgtgctt 1578

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: gb_est1.*

2: gb_est2.*

3: gb_hc.*

4: gb_est3.*

5: gb_est4.*

6: gb_est5.*

7: gb_est6.*

8: gb_gsl1.*

9: gb_gsl2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	407.6	25.8	1522	4	BM320865
2	397.6	25.2	2491	8	BH770842
3	386.8	24.5	1626	3	CR685174
4	385.2	24.4	1643	3	CR728296
5	384.2	24.3	1505	4	BM321196
6	382	24.2	1623	3	CR679338
7	377.4	23.9	1613	3	CR698859
8	377.4	23.9	1615	3	CR650019
9	377	23.9	1637	3	CR728153
10	369.6	23.4	1551	3	CR685932
11	366.8	23.2	1641	3	CR672414
12	364.6	23.1	1582	3	CR656482
13	341.8	21.7	2021	8	BH770828
14	328	20.8	1551	3	CR658829
15	327	20.7	1407	3	CR733357
16	326.4	20.7	1476	3	CR726536
17	326.4	20.7	1563	3	CR657796
18	326.4	20.7	1587	3	CR675445
19	326.4	20.7	1592	3	CR664987
20	326.4	20.7	1592	3	CR731975
21	325.6	20.6	1588	3	CR731270
22	325.2	20.6	1566	3	CR729213
23	324.8	20.6	1518	3	CR675234
24	324.8	20.6	1536	3	CR670482

25	324.8	20.6	1538	3	CR669496	Tetraodon
26	324.8	20.6	1543	3	CR659851	Tetraodon
27	324.8	20.6	1547	3	CR678580	Tetraodon
28	324.8	20.6	1549	3	CR676296	Tetraodon
29	324.8	20.6	1551	3	CR658992	Tetraodon
30	324.8	20.6	1554	3	CR650486	Tetraodon
31	324.8	20.6	1557	3	CR685781	Tetraodon
32	324.8	20.6	1568	3	CR664375	Tetraodon
33	324.8	20.6	1583	3	CR676706	Tetraodon
34	324.8	20.6	1596	3	CR724795	Tetraodon
35	324.8	20.6	1598	3	CR673577	Tetraodon
36	324.8	20.6	1598	3	CR729481	Tetraodon
37	324.8	20.6	1606	3	CR723707	Tetraodon
38	324.8	20.6	1611	3	CR726056	Tetraodon
39	324.8	20.6	1613	3	CR728206	Tetraodon
40	323.6	20.5	1534	3	CR646460	Tetraodon
41	323.2	20.5	1520	3	CR653599	Tetraodon
42	323.2	20.5	1533	3	CR649337	Tetraodon
43	323.2	20.5	1541	3	CR656841	Tetraodon
44	323.2	20.5	1541	3	CR678734	Tetraodon
45	323.2	20.5	1547	3	CR723775	Tetraodon

ALIGNMENTS

RESULT 1
BM320865 1522 bp mRNA linear EST 03-JAN-2002
LOCUS rocketfeller.0.878 Mastigamoeba balamuthi lambda ZAP II library
DEFINITION Mastigamoeba balamuthi cDNA similar to enolase (EC 4.2.1.1), mRNA sequence.
ACCESSION BM320865
VERSION BM320865.1 GI:18055271
KEYWORDS EST.
SOURCE Mastigamoeba balamuthi
ORGANISM Mastigamoeba balamuthi
REFERENCE 1 (bases 1 to 1522)
AUTHORS Eukaryota; Pelobiontida; Mastigamoebidae; Mastigamoeba.
Bapteste,E., Brinkmann,H., Lee,J.A., Moore,D.V., Sensen,C.W., Gordon,P., Durufle,L., Gaasterland,P., Lopez,P., Muller,M. and Philippe,H.
TITLE The analysis of 100 genes supports the grouping of three highly divergent amoebae: Dictyostelium, Entamoeba, and Mastigamoeba
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (3), 1414-1419 (2002)
MEDLINE 21819461
PUBMED 11830664
COMMENT Contact: Muller Miklos
Laboratory of Biochemical Parasitology
The Rockefeller University
1230 York Avenue, New York, NY 10021, USA
Email: mmuller@rockvax.rockefeller.edu
Insert Length: 1522 Std Error: 0.00
POLYA=Yes.

FEATURES
source Location/Qualifiers
1..1522
/organism="Mastigamoeba balamuthi"
/mol_type="mRNA"
/db_xref="taxon:108607"
/clone_lib="Mastigamoeba balamuthi lambda ZAP II Library"
/note="syn: Phreatamoeba balamuthi"

ORIGIN
Query Match 25.8%; Score 407.6; DB 4; Length 1522;
Best Local Similarity 60.4%; Pred. No. 1.3e-98;
Matches 787; Conservative 0; Mismatches 464; Indels 51; Gaps 5;
Qy 143 AGCCACAGTGGTGAATCATGACGATATTCGTCGGAATTCGACTCCCGCGTA 202
Db 30 AAGCAACAATGTGACCATCAAGAGCTCTTCGCCCGGAGATCTCGTGGCA 89
Qy 203 ACCCAACCGTCGAGGAGAGGTTTTCTCGATGACGGTCCACGGTGTGCGAGGTGTC 262

```
Db 90 ACCCCACCGTCGAGGTTGACCTGACCAACGAGAGGCGCTCTTCCG---CTCGCCGCTCC 146
QY 263 CATCCGCGCATCACCGGCGTCCAGAGGCTCATGAGCTGCGTGAGCGTGGCGA---TC 319
Db 147 CTTCTGTGCTGACTGGGATCTATGAGGCTGCGGAGCTCCGTGACGGCGACAAGTCGC 206
QY 320 GCTACCTTGGCAAGGGCGTTTGAAGGCACTTGAAGAAAGTCAACGAGAAATCGGCGAG 379
Db 207 GCTACCTTGGCAAGGGGTGCTGAAGGCGCTGAGAACGTCACAAAGATCCTGGCGCCA 266
QY 380 AGCTCGTGGCGCTAGAGCTGACGATCAGCGCTCATCGAGCAAGCAATGATCAAGCTTG 439
Db 267 AGCTGATCGGCTCGAGCTCACCAAGCAGGCGGAGATCGACAGGCTCATGCTCAAGATCG 326
QY 440 ATGSCACCGCAACAAGTCCCGCTGGGTGCAACGCAATCCTTGGTGTGTTTCCATGGCTG 499
Db 327 ACGGCACTGAGAACAGACCACTGGGCGGCAAGCCATCCTCGGCTGCTCGATGTGG 386
QY 500 TTGCAAAAGGCTGTGCTGATTCGCGAGGCTCCACATGTTCCGCTACATCGGT----- 552
Db 387 TCTGCCGTGGCGCGCGGCTTACAGGCGCTGCGCCCTCTACCGCTACATCGCGAGCTGT 446
QY 553 --GGACCAACGACACGTTCTTCCAGTTCCAATGATGAACATCATCAACGCTGGGCTC 610
Db 447 CGGCAACAAGTCCGCTATGCTCTTGCCTCTGCTTCAAGCTCATCAACGCGGCTGAGC 506
QY 611 ACGTGTACTCGGTGTGAGTTCAGGAATTCATGATCGCTCCAAATCGGTGCGAGACCT 670
Db 507 ACGGGGCAACAAGCTCGCCATCGAGGATTCATGATTCGCCCAACGCGGCGACCTCGT 566
QY 671 TCTCTGAGGCTCTCCGCAACGCGCGGAGGCTTACCAAGCACTGAAGTCGCTCATCAAGG 730
Db 567 TCCACGAGGCGCTCGCATGCGCCGCGAGAGTACCAACCTGAAGCTGCTCATCAAGA 626
QY 731 AABAGGGC-----CTGTCCACCGGACTTGGCGATGAGGGGCTTCGCTCTTCCGTCG 784
Db 627 AGAGGTACGGATGGACGCGACCAACGTCGCGGACGAGGGTGTTCGCCCCCAACATCC 686
QY 785 GCTCACCGCTGAGGCTCTTGACCTTATCGTTGAGGCAATCGAAGGCTGGCTTACCC 844
Db 687 AGGCCAACCAACGAGGCGCTCGAGCTCATCGTTGAGGCCATCAAGCAGCGCGCTACACTG 746
QY 845 CAGGCAAGGACATCGCTCTGCTCTGAGCTGTGCTCTCTGAGTCTTCAAGACAGGCA 904
Db 747 GCAAGATCGAGATCGGCATGGACGTTCGCGCTCGTCTGCGGACGCCAAGAGTGA 806
QY 905 CTTACCACTTCGAAGTGGCCAGCACTCCGACGCTGAGA----- 943
Db 807 AGTAGCACTCGGCTTCAAGTTCGCCGCGACAGAAGACCCCGACATGCTGCTCTCGG 866
QY 944 -----TGGCAACGTTTACGCTGAGCTCGTTGACGCGTACCCAAATGCTTCCATCG 994
Db 867 GCGAGGCGCTGATCAAGCTTACGAGGAGTGGACCTCGAAGTACCGATCTGCTCGATCG 926
QY 995 AGGACCACTGCAGGAAGATGACTGGGAGGTTACACCAACCTCACCAACCATCGCG 1054
Db 927 AGGACCCCTTCGACAGGACGACTGGGCCA CTTACACCAAGTTTACCAGGCTGATCGGCA 986
QY 1055 ACAAGGTTTCAGATCGTTGGCGACGACTTCTTCGTCAACCAACCTCGAGCGCTGAAGGAGG 1114
Db 987 ACAGATCCAGATCGTCGGCGAGGACTCTGCTGCTCAACCAACCCCAAGCGCATCGTCGAGG 1046
QY 1115 GCATCGCTAAGAGGCTGCGCAACTCCATCTCGGTTAAGGTGAACAGATCGGTACCTCA 1174
Db 1047 CCGCAACAAGAGGCTGCAACGCGCTGCTGCTCAAGCTCAACAGATCGGCTCGGTCA 1106
QY 1175 CGGAGCCTTCGACCTGTGACATGCTCACCGCGCAGGCTTACACCTCCATGATGTC 1234
Db 1107 GCGAGGCGTTCGAGGCGTGCCTGTCGGCGGAGGTCAACTGGGGCGTCAATGCTGTCG 1166
QY 1235 ACCGTTCCGGTGAGACCGAGGACACCAATTCCTGACCTCGCAGTTCGACTCAACTGTG 1294
```

```
Db 1167 ACCGTCGGCGAGACTGAGGACGCGTTTCATCGCCGACCTCGTCTGGTCTGGGCTGG 1226
QY 1295 GCCAGATCAAGACTCGTGTCTCCAGCAGCTTCCGACCGTGTCCGAAAGTACACCGACTTC 1354
Db 1227 GCCAGATCAAGACTCGTGTCCCTCGCTCGGAGCGCTCGCAAGTACAAACCGACTCC 1286
QY 1355 TCCGATCGAGCAGCTGCTTGGCGACGCGCGGCTCTACGCG 1396
Db 1287 TCCGATCGAGGAGGCTCGGCGCAACGCCCACTACGCCG 1328
```

RESULT 2
BH770842/c

LOCUS BH770842/c 2491 bp DNA linear GSS 01-MAY-2002
DEFINITION LJMtag581 MG1363 Random Sequence Tag Library Lactococcus lactis
subsp. cremoris genomic, Genomic survey sequence.

ACCESSION BH770842

VERSION BH770842.1 GI:20373799

KEYWORDS GSS

SOURCE Lactococcus lactis subsp. cremoris

ORGANISM Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;

Lactococcus.

REFERENCE 1 (bases 1 to 2491)

AUTHORS Bolotin,A., Ehrlich,S.D. and Sorokin,A.

TITLE Studies of genomes of dairy bacteria Lactococcus lactis

JOURNAL Sci. Aliments (2002) In press

COMMENT Contact: Sorokin A

Genetique Microbienne

INRA

CRJ INRA, Domaine de Vilvert, 78352 Jouy en Josas cedex, France

Tel: 33 1 34 65 25 16

Fax: 33 1 34 65 25 21

Email: sorokine@jouy.inra.fr

Best homologue in strain IL1403 is enoA (100%)

Class: shotgun

High quality sequence start: 30

High quality sequence stop: 2463.

Location/Qualifiers

source

1. 2491

/organism="Lactococcus lactis subsp. cremoris"

/mol_type="genomic DNA"

/strain="MG1363"

/sub_species="cremoris"

/db_xref="taxon:1359"

/clone_lib="MG1363 Random Sequence Tag Library"

/notes="Vector: pSGM02; Site1: Smal; Library of

chromosomal fragments of L.lactis strain MG1363 was

prepared by partial AluI digestion or by sonication."

ORIGIN

Query Match 25.2%; Score 397.6; DB 8; Length 2491;

Best Local Similarity 60.4%; Pred. No. 7.5e-36;

Matches 721; Conservative 0; Mismatches 449; Indels 24; Gaps 3;

QY 259 GTTCCATCCGCGCATCCACCGGCTCCACGAGGCTCATGAGCTCGGTGCGTGGCGCA- 317

Db 2485 GTACCTTCAGTGTCTTCTACTGTGTGAACACGAGCGGTGAACCTCGTATGGCGACAA 2426

QY 318 --TGGTACCTGGGCAAGGCGGTTTGAAGCGAGTTGAAAACGTCAACGAAGAATCGGC 375

Db 2425 TCTCGCTACAAACGACGCTTGGTACTCAAAAAGCTGTTGACAACGTAACCATCATCGCT 2366

QY 376 GAGCAGCTCGCTCGCTTAGAGGCTGACGATCAGCGCTCATCGACGAAGCAATGATCAAG 435

Db 2365 GAAGCTATCATCGGTTATGAATTACTGACCAACAGCTATTGACCGTGAATGATCGCT 2306

QY 436 CTTGATGGCCCGCCAAACAAGTCCCGCTGGGTGCAACGCAATCTCTGGTGTTCATG 495

Db 2305 CTTGACGCTACTGAAAACAAGGTAATTTGGAGCTAACGCTATTCTTGTGTTTCTATC 2246

QY 496 GCTGTTGCAAGGCTGCTGCTGATTCGCGAGGCTCCCACTGTTCCCGCTACATCGGTGA 555

```
Db 2245 GCTGCTGCTGCTGCTGCTGATGAACCTTGGTGTTCACCTTTAAACTACCTTGGCGGA 2186
Qy 556 CCAAAACGACACAGCTTCTTCAGTTCCTCAATGATGAACATCATCAACGGTGGCGCTCACGCT 615
Db 2185 TTCAACGCTAAAGTATTGCCAACTCCATGATGAACATCATCATGTTGGTTCATCTCA 2126
Qy 616 GACTCCGGTGTGACGTTAGGAAATTCATGATGCTCCATTCGGTGGAGACGCTTCTCT 675
Db 2125 GACGCCCTTATCGCTTTCAGAAATTCATGATGCTGACGATGCTGTCACCTACATTCAAA 2066
Qy 676 GAGGCTCTCCGCAACGGCGGAGGCTTACCACGACTGAGTCCGTCATCAAGGAAAG 735
Db 2065 GAAGCGCTTCGTTGGGTGCTGAATCTTCACGCTCTTAAAGAAATTCCTAAAGCTCGT 2006
Qy 736 GGCCTGTCCACCGGACTTGGCGATGAGGGGGCTTCGCTCCTCCGTCGGCTCCACCGT 795
Db 2005 GGACTTGAACAGCTGCGGTGACGAAGTGGATTGCTCTTAATTCGACGGAACCTGAA 1946
Qy 796 GAGGCTCTTGACCTTATCGTTGAGGCAATCGAAGGCTGGCTTCAACCCAGGCAAGGAC 855
Db 1945 GACGGTGTAGAAACTATCTTAAAGCAATCGAAGCAGCTGGTTACAAAGCTGGTGAAGAT 1886
Qy 856 ATCGCT---CTTGCTCTGAGCTTCTTCTCTGAGTTCCTCAAGGACGGCACCTACCAC 912
Db 1885 GGGCTTATGATCGGTTTCGACTGTGCATCATCAGAAATTCACGAAACCGTGTTCAGC 1826
Qy 913 TTC-----GAAAGTGGCGAGCACTCCGCGCTGAGATGGCAACGTT 954
Db 1825 TACACTAAATTCGAAGGTGAAGGGGTAAAGAACTTTCAGCTTCTGAACAGTTGACTAC 1766
Qy 955 TAGCTGAGCTCGTTGACGGGTACCCAAATCGTCTCCATCGAGGACCCACTGCGAGGAAT 1014
Db 1765 CTTGAAGAACTCGTTTCTAAATACCCAAATCATCACTATTGAAGATGCTATGGACGAAAC 1706
Qy 1015 GACTGGAGGTTTACACCACTCCAGCAACCATCGCGACAGGTTCAGATCGTTGGC 1074
Db 1705 GACTGGGATGGATGGAAATCTCTTACTGAAAGCTTCTGGTAAAGAAAGTTCAACTCGTTGGT 1646
Qy 1075 GACGACTTCTTCTGTCACCAACCTCGAGCGCTGAAGGAGGCGATCGCTAAGAGGCTGCC 1134
Db 1645 GACGACTTCTTCTGTTACAAACACTAAATACCTTGAAGTGGTATCCGTGAATGCTTCA 1586
Qy 1135 AACTCCATCTCGTTAAGGTGAACAGATCGGTACCTCAACCGAGACCTTCGACGCTGTC 1194
Db 1585 AACGCTATCTGATCAAAATTAACCAATCGGTACTTTGACAGAAACTTTTCGAAGCTATT 1526
Qy 1195 GACATGCTCACCGGAGGCTACACCTCCATGATGCCACCGTTCGGGTGAGACCGAG 1254
Db 1525 GAAATGGCTAAAGAAAGCTGGTTTACAGCAATCGTATCTCACCGTTTCAGGTGAACCTGAA 1466
Qy 1255 GACACCACTTGTGACCTCGCAGTTGCACTCAACTGTGGCCAGATCAAGACTGGTGCT 1314
Db 1465 GATTCAACATCTCAGACATCGCTGTTGCAACTAACGCTGGTCAATCAAACTGGTTCA 1406
Qy 1315 CCAGCAGCTTCGACCGGTGTGCAAGGTACAAACAGCTTCTCCGCACTCGAGCAGCTGCTT 1374
Db 1405 CTTTCAGCTACAGACCGGTATGGCTTAAATACAACTTGTCTGTTTGAAGCACTTGG 1346
Qy 1375 GCGACCGCGGCTCTAGGAGGTGCGAGCGCATTCACGCTTTCAGGGGTAA 1428
Db 1345 GCTGAAGTTGCTCAATAAAGGCTTAAAGCACTTCTACACCTTAAAGAAATTA 1292
```

```
RESULT 3
CR685174
LOCUS
DEFINITION Tetraodon nigroviridis full-length cdna.
ACCESSION CR685174
VERSION CR685174.1 GI:51183081
KEYWORDS HTC; cdna; full-length; Tetraodon nigroviridis.
SOURCE Tetraodon nigroviridis
ORGANISM Tetraodon nigroviridis
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
Actinopterygii; Neopterygii; Teleostei; Buteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodontidae; Tetraodon.
1 (bases 1 to 1626)
Genoscope.
Direct Submission
Submitted (10-AUG-2004) Genoscope - Centre National de Sequencage -
: 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
(E-mail : segref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
The sequences are based on single pass reads.
More information available at
http://www.genoscope.cns.fr/tetraodon.
Location/Qualifiers
    source          1..1626
                    /organism="Tetraodon nigroviridis"
                    /mol_type="mRNA"
                    /db_xref="taxon:99883"
                    /tissue_type="Eyes"

ORIGIN
Query Match      24.5%; Score 386.8; DB 3; Length 1626;
Best Local Similarity 60.9%; Pred. No. 5.5e-93;
Matches 798; Conservative 0; Mismatches 452; Indels 60; Gaps 8;

Qy 160 ATCATGCACGATATTCGCTCGCGAAATTCGACATCCCGGTAACCCAAACCGTCGAGGCA 219
Db 110 ATCTGAAGATCCACGCTCGGAAATATTCGATCCCGCGCAACCCACCGTCGAGGTC 169
Qy 220 GAGGTTTCTGGATGAGGTTCCACGCTGTCGAGGTGTTCCATCGGCGCATCCACC 279
Db 170 GACCTGTACACCAAGAAAGGTCTGTTGAGAGTGTG---GTCCCGAGCGGCGCTCCACC 226
Qy 280 GGGCTCCACGAGGCTCATGAGCTGCGTGCGGTGGCGA---TCGCTACTCTGGCAAGGCG 336
Db 227 GGCATCTACGAGCGCTTGGAGCTCCGCGACAAACGAAACCCGCTATATGGGCAAGGT 286
Qy 337 GTTTTGAAGCAGTTGAAACGTCACGAAAGAAATCGGCGACGAGTCTGCTGGCTAGAG 396
Db 287 GTCTTAAAGCTGTGAGATATCAATAAATAATTCACCTGCACTGGTTGGCAAGGAC 346
Qy 397 GCTGACGATCAGGCGC-----TCATCGACGAAGCAATGATCAAGCTTGTGGCACCGCC 450
Db 347 GTCAACGTTCTGGACACGAGGCAAAATCGCAAGCTGATGCTGGAGATGGACGCGACCGAC 406
Qy 451 AACAAAGTCCCGCTGGGTGCAACGCAATCTTGTGTTTCCATGGCTGTTGCAAGGCT 510
Db 407 AACAAATCTAAGTTCGGGGCAACGCCATCTCTGGCGTGTCCCTGGCTGTGTGCAAGGCT 466
Qy 511 GCTGCTGATTCGCGAGGCTCCCACTGTTCCGCTPACATCGGTGGACCAAAACGC----- 563
Db 467 GGTGACGACAGAGGAGGCGTGCACCTCTACCGCACATCGCCGACCTGGCGGCAACCCC 526
Qy 564 --ACAGTTCCTTCAGTTTCCAATGATGAACATCATCAACGGTGGCGCTCAGCTGACTCC 621
Db 527 CAAGTCACTCTCCCGTTCGCGCTTTTCAACGTCATCAACCGCGGCTCCACGCGAGGAAAC 586
Qy 622 GGTGTTGACGCTTCAGGAATTCATGATCGCTCCAATCGTGCAGAGACCTCTCTGAGGCT 681
Db 587 AAGTGGCCATGACGAGGTTCAATGATCTCTCGGTGGAGCCAGCAGCTTCAAGAGGCGC 646
Qy 682 CTCGCGAACCGCGGAGGTTTACACGCACTGAAGTCCGTCATCAAGGAAAGGCGC--- 738
Db 647 ATGCGCATCGGCGCGAAGTCTACCAACAACTGAAGAAACGTCATCAAGGAGGATACGGC 706
Qy 739 ---CTGTTCACCGGACTTGGGATGAGGGGCTTCGCTCTCTCCGTGGCTCCACCGT 795
Db 707 AAGGACGCCCAACGCTGGGAGACGAGGAGGCTTTGCCCGCCCAACATCTCTGGAGAACAA 766
Qy 796 GAGGCTCTTGACCTTATCGTTGAGGCAATCGAAGGCTGCTTCAACCCAGGCAAGGAC 855
Db 767 GAAGCTCTGGAGCTGTGAAGAACGCCATCGTAAGCGCGGCTACAC-----CGACAAG 820
Qy 856 ATCGCTCTTCTCTGGACGTTGCTTCTCTGAGTTCTTCAAGGACGCGACCTACCACTTC 915
```

```

Db      821  ATTGTGATCGGCATGACGCTGGCTGCTCTAGTTCTTACAAGGGCGGCAAGTACGACCTG 880
Qy      916  GAA-----GGTGGCCAGCACTCCGCACTCAGATGGAACAGTT 954
Db      881  GACTTCAAGTCTCCGACGACCGCGCGCTACATCTCTCCGAGCAGCTGGCGACCTC 940
Qy      955  TACGCTGAGCTCGTTGACGGTACCCAAATCGTCTCCATCGAGACCCACTGCGAGGAAT 1014
Db      941  TACAGGGGCTTCGTCAAGATTACCCAGTGGTTCATCGAGACCCCTTTGACCAAGGAC 1000
Qy      1015  GACTGGGAGGTTACACCACTTACCGCAACCATCGGCGACAAGTTCAGATCGTTGGC 1074
Db      1001  GACTGGGAGGCTGGACCAATTTACAGCCAGCACCAAGC-----ATTAGGTGGTGGC 1054
Qy      1075  GACGACTTCTTCGTCAACCACTTCAGCGCTGAAGGGGCTCGTAAAGAGGCTGCC 1134
Db      1055  GACGACTTCAAGTCAACCGACCCCAAGCATCGCCNAGGCTGGCCCAAGTCTCTGC 1114
Qy      1135  AACTCCATCTCGTTAAGGTGAACCAAGATCGGTACCTCAACGAGCTTCGACGCTGTC 1194
Db      1115  AACTGCTGCTGTCAAGTCAACAGATCGGCTCGGTCAACGAGTCCCTGCGAGGATGC 1174
Qy      1195  GACATGCTCACCAGGAGGCTACACCTCCATGATGTCGCCACCGTTCGGTGAGACGAG 1254
Db      1175  AAGATGGCCAGAGCAACGGCTGGGGCTGATGTCAGCCATCGCTCCGGGGAGACGGAG 1234
Qy      1255  GACACCACTTGTGACCTCGCAGTTGCACTCAACTGTGTGGCCAGATCAAGACTGGTGT 1314
Db      1235  GACACCTTCACTGCTGACCTGGTGGTCTGTGACACCGGACAGATCAAGACGGGGCT 1294
Qy      1315  CCAGCAGTTCGACCGCTGTGCGAAAGTACAACAGATTCCTCGCATCGAGCAGCTGCTT 1374
Db      1295  CCGTGGCGCTCCGAGCGCTTGGCCAAAGTACAACAGCTGTCTCGGATCGAGGAAGCTG 1354
Qy      1375  GCGACCGCGCTTACGAGGCTCGCAGCGCATCCACGCTTTCAGGG 1424
Db      1355  GCGACCGCGCTTCCGCGGCAAGAACTTCAGGCACCCCATCTGAGCG 1404

RESULT 4
LOCUS   CR726296 1643 bp mRNA linear HTC 19-AUG-2004
DEFINITION Tetraodon nigroviridis full-length cDNA.
ACCESSION CR726296
VERSION   CR726296.1 GI:51224547
KEYWORDS HTC; cDNA; full-length; Tetraodon nigroviridis.
SOURCE   Tetraodon nigroviridis
ORGANISM Tetraodon nigroviridis
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
           Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
           Tetraodontidae; Tetraodontinae; Tetraodon.
           1 (bases 1 to 1643)
REFERENCE Genoscope.
AUTHORS Direct Submission
TITLE Submitted (10-AUG-2004) Genoscope - Centre National de Sequencage -
JOURNAL : 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
          (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT The sequences are based on single pass reads.
          More information available at
          http://www.genoscope.cns.fr/tetraodon.
FEATURES
   source
       1..1643
           /organism="Tetraodon nigroviridis"
           /mol_type="mRNA"
           /db_xref="taxon:99883"
           /tissue_type="fish"

ORIGIN
Query Match 24.4%; Score 385.2; DB 3; Length 1643;
Best Local Similarity 60.8%; Pred. No. 1.5e-92;
Matches 797; Conservative 0; Mismatches 453; Indels 60; Gaps 8;

```

```

Qy      160  ATCATGCAGTATTCTGCTCGCGAAATTTCTGACTCCCGGGTAAACCCACCGTCGAGCA 219
Db      113  ATCTTGAAGATCCACGCTCGGAAATATTCGACTCCCGGCAACCCACCGTGAGGTTC 172
Qy      220  GAGGTTTTCTCGATGACGGTTCCACCGTGTGCGAGGTGTTCATTCGGCGCATCCACC 279
Db      173  GACCTGTACCAAGAAAGTCTGTTCAGAGTGG---GTCCCGAGCGCGCTCCACC 229
Qy      280  GCGCTCCAGAGGCTCATGAGTGTGCTGACGCTGGCGA---TCGTTACCTGGGCAAGGC 336
Db      230  GGCATCTACGAGGCTCGAGCTCGCGACAACCAAAACCCGCTTACATGCGCAAGGT 289
Qy      337  GTTTTGAAGGAGTTGAACACGTCACAGNAGAAATCGGCGAGAGCTCGCTGCGCTTAGAG 396
Db      290  GTCTCTAAAGCTGTTGAGAAATATCAATAAATAATTTGACCTGSCACTGGTTGGCAAGGAC 349
Qy      397  GCTGACGATCAGCGC-----TCATCGACGAAGCAATGATCAAGCTTGTATGGCACCGC 450
Db      350  GTCACGTTCTGGACCGCAAGCAAAATCGACAGCTGATGCTGGAGATGACGCGACCGAC 409
Qy      451  AACAAAGTCCCGCTGGGTGCAAAACGCAATCTTGGTGTTCATGCTGCTGTTGCAAGGCT 510
Db      410  AACAAATCTAAGTTGCGGGCCAAACGCAATCTTGGGCGTGTCCCTGGCTGTGTGCAAGGCT 469
Qy      511  GCTGCTGATTCGCGAGGCTCCCACTGTTCCGCTACATCGGTGGACCAACGC-----563
Db      470  GGTGACGACGAGAAAGGGCGTGCCTCTACCGCCATCGCCGACCTGGCGGCAACCCC 529
Qy      564  --ACAGTTCTTTCAGTTCCAATGATGAACATCATCAACGGTGGGCTCACGCTGACTCC 621
Db      530  CAAAGTCATCTCCCGTTCCCGCTTTCACAGTTCATCAACGGGGCTCCACGCGAGCAAC 589
Qy      622  GGTGTTGACGTTACAGAAATCATATGATCGCTCCAAATCGGTGCGAGACCTTCTCTGAGGCT 681
Db      590  AAGCTGGCCATGCAGGAGTTTATGATCTCTGCGGTGGAGCCAGCAGCTTCAAAGAGGCG 649
Qy      682  CTCGCAACGGCGCGAGGCTTACCAACGCACTGAAGTCGTCATCAAGGAAAGAGGC---738
Db      650  ATGCGCATCGGCGCCGAGTCTACCAACACCTGAAGAACGTCAATCAAGAGAAATAGGCG 709
Qy      739  ---CTGTCCACCGGATTTGGCGATGAGGGCGGCTTCGCTCTTCCGTGGCTCCACCGCT 795
Db      710  AAGGATGCCACCAACCTGGGAGACGAGGAGGCTTTGCCCCCAACATCTTGGAGAAACAAG 769
Qy      796  GAGGCTCTTGACCTTATCGTTGAGGCAATCGAGAAAGCTGGCTTACCCCGCAGCAAGGAC 855
Db      770  GAAGCTCTGGAGCTGCTGAAGAACGCGCATCGCTTAAGCGCGGCTTACAC-----CGACAAG 823
Qy      856  ATCGCTCTTGTCTTGACGCTTGTCTCTGAGTTCTTCAAGGACGCGACCTACCACTTC 915
Db      824  ATTGTGATCGGATGAGCGTGGCTGCTCGAGTTCTCAAGGGCGGCAAGTACGACCTG 883
Qy      916  GAA-----GGTGGCGACACTTCGCGACGTTCGCGACGTGAGATGGCAAAAGCTT 954
Db      884  GACTTCAAGTCTCCCGACGACCCCGCGCTACATTTCTCCGAGCAGCTGGCGCGACTC 943
Qy      955  TAGCTGAGCTGTTGACGGGTACCCAAATCGTCTCCATCGAGACCCACTGCGAGGAGAT 1014
Db      944  TACAGGGGCTTCGTCAAGATTACCCAGTGGTGTTCATCGAGACCCCTTTTGACCAAGGAC 1003
Qy      1015  GACTGGGAGGTTTACACCAACCTCACCGCAACCATCGGCGACAAGGTTTCAGATCGTTGGC 1074
Db      1004  GACTGGGAGGCTGGACCAATTTCAAGCAGCAGCACCAAGC-----ATTAGGTGGTGGC 1057
Qy      1075  GACGACTTCTTCGTCAACCAACCTGAGCGCTTGAAGAGGGGCTATCGTGAAGAGGCTGCC 1134
Db      1058  GACGACCTCAGGTCACCAACCCCAACGCAATCGCCAAAGGTTGTGGCCCAAGTCTCTGC 1117
Qy      1135  AACTCCATCTGTTAAGGTGAACCGATCGGTACCTCAACCGAGACCTTCGAGCGCTGTC 1194
Db      1118  AACTGCTGCTGCTCAAAAGTCAACAGATCGGCTCGGTCCGCTACCGAGTCCCTGCGAGGATGC 1177

```

QY	1195	GACATGCTACCGCGGAGGCTACACCTTCATGATGTGCCACCGTTCCGGTGTGAGACCGAG	1254
Db <th>1178</th> <th>AAGATGCCCCAGAGCAATGGCTGGGGCGTGTATGTGTACGCCATCGCTCCGGGAGACGCGAG</th> <th>1237</th>	1178	AAGATGCCCCAGAGCAATGGCTGGGGCGTGTATGTGTACGCCATCGCTCCGGGAGACGCGAG	1237
QY <th>1255</th> <th>GACACCACTTGTGTGACCTCGCAGTTGTGACCTCAACTGTGTGCCAGATCAAGACTGTGTGCT</th> <th>1314</th>	1255	GACACCACTTGTGTGACCTCGCAGTTGTGACCTCAACTGTGTGCCAGATCAAGACTGTGTGCT	1314
Db <th>1238</th> <th>GACACTTTCATCGTGTGACCTGGTGTGCTGTGCGCACCGGACAGATCAAGACGGCGCT</th> <th>1297</th>	1238	GACACTTTCATCGTGTGACCTGGTGTGCTGTGCGCACCGGACAGATCAAGACGGCGCT	1297
QY <th>1315</th> <th>CCAGCAGCTTCCGACCGTGTGCGAAAGTACAAACAGCTTTCCTCCGCATCGAGCAGCTGCTT</th> <th>1374</th>	1315	CCAGCAGCTTCCGACCGTGTGCGAAAGTACAAACAGCTTTCCTCCGCATCGAGCAGCTGCTT	1374
Db <th>1298</th> <th>CCGTGCGCTCCGAGCGCTTGGCCAAAGTACAAACAGCTGTCTCCGATCGAGGAAGAGCTG</th> <th>1357</th>	1298	CCGTGCGCTCCGAGCGCTTGGCCAAAGTACAAACAGCTGTCTCCGATCGAGGAAGAGCTG	1357
QY <th>1375</th> <th>GGCGACCGCGCGTCTAGCGAGTGTGCGAGCGCATTTCCAGCTTTTCAGGG</th> <th>1424</th>	1375	GGCGACCGCGCGTCTAGCGAGTGTGCGAGCGCATTTCCAGCTTTTCAGGG	1424
Db <th>1358</th> <th>GGCGACGAGCGCGCTTCCGCGGCAAGAACTTCAGGCACCCCATCTGAGCG</th> <th>1407</th>	1358	GGCGACGAGCGCGCTTCCGCGGCAAGAACTTCAGGCACCCCATCTGAGCG	1407
RESULT 5			
BM3211196	1505 bp mRNA linear EST 03-JAN-2002		
LOCUS	rockefeller.0.899 Mastigamoeba balamuthi lambda ZAP II Library		
DEFINITION	Mastigamoeba balamuthi cDNA similar to enolase (EC 4.2.1.1), mRNA		
sequence.			
ACCESSION	BM3211196		
VERSION	BM3211196.1 GI:18055602		
KEYWORDS	EST.		
SOURCE	Mastigamoeba balamuthi		
ORGANISM	Mastigamoeba balamuthi		
REFERENCE	1 (bases 1 to 1505)		
AUTHORS	Baptiste,E., Brinkmann,H., Lee,J.A., Moore,D.V., Sensen,C.W., Gordon,P., Durufle,L., Gaasterland,T., Lopez,P., Muller,M. and Philippe,H.		
TITLE	The analysis of 100 genes supports the grouping of three highly divergent amoebae: Dictyostelium, Entamoeba, and Mastigamoeba		
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (3), 1414-1419 (2002)		
MEDLINE	21819461		
PUBMED	11830664		
COMMENT	Contact: Muller Miklos Laboratory of Biochemical Parasitology The Rockefeller University 1230 York Avenue, New York, NY 10021, USA Email: mmuller@rockvax.rockefeller.edu Insert Length: 1505 Std Error: 0.00 POLYA=Yes.		
FEATURES			
source	Location/Qualifiers		
	1..1505		
	/organism="Mastigamoeba balamuthi"		
	/mol_type="mRNA"		
	/strain="ATCC 30984"		
	/db_xref="taxon:108607"		
	/clone_lib="Mastigamoeba balamuthi lambda ZAP II Library"		
	/note="syn: Phreatamoeba balamuthi"		
ORIGIN			
Query Match	24.3%	Score 384.2; DB 4; Length 1505;	
Best Local Similarity	60.08;	Pred. No. 2.7e-92;	
Matches	782; Conservative	0; Mismatches 468; Indels 53; Gaps 7;	
QY <th>143</th> <th>AGGCCACAGTGGCTGAAATCATGCACGTATTTCGTCTCGCGAAATTTCTCGACTCCCGCGGTA</th> <th>202</th>	143	AGGCCACAGTGGCTGAAATCATGCACGTATTTCGTCTCGCGAAATTTCTCGACTCCCGCGGTA	202
Db <th>27</th> <th>AAGCAACAATGTCGACCATCAAGAGCGTCTTCGCGCGCGAGANCTTGGACTCTCGTGCGCA</th> <th>86</th>	27	AAGCAACAATGTCGACCATCAAGAGCGTCTTCGCGCGCGAGANCTTGGACTCTCGTGCGCA	86
QY <th>203</th> <th>ACCCACCGTCTGAGGACAGGTTTTTCTTGATGACGTTTCCACGTTGTCGAGTGTTCGAGTTTC</th> <th>262</th>	203	ACCCACCGTCTGAGGACAGGTTTTTCTTGATGACGTTTCCACGTTGTCGAGTGTTCGAGTTTC	262
Db <th>87</th> <th>ACCCACCGTCTGAGGTTGACTGACCAACCGAGAGGGCGCTTTTCG---CTCGGCGCTCC</th> <th>143</th>	87	ACCCACCGTCTGAGGTTGACTGACCAACCGAGAGGGCGCTTTTCG---CTCGGCGCTCC	143
QY <th>263</th> <th>CATCCGCGCATCCACCGCGTCTCACGAGGCTCATGAGCTGCGTGACGTTGGCA---TC</th> <th>319</th>	263	CATCCGCGCATCCACCGCGTCTCACGAGGCTCATGAGCTGCGTGACGTTGGCA---TC	319
Db <th>144</th> <th>CTTCTGTTGCTTCGACTGGGATCTATGAGGCTCGGAGCTTCGTCGCGCAAGTCGC</th> <th>203</th>	144	CTTCTGTTGCTTCGACTGGGATCTATGAGGCTCGGAGCTTCGTCGCGCAAGTCGC	203
QY <th>320</th> <th>GCTACTTGGCAAGGGCGTTTGTGAAGCGAGTTGAAACGTCACCAAGAAATCGGCGACG</th> <th>379</th>	320	GCTACTTGGCAAGGGCGTTTGTGAAGCGAGTTGAAACGTCACCAAGAAATCGGCGACG	379

```

RESULT 6
CR679338
LOCUS          CR679338          1623 bp      mRNA      linear      HTC 19-AUG-2004
DEFINITION    Tetraodon nigroviridis full-length cDNA.
ACCESSION     CR679338
VERSION       CR679338.1 GI:51177245
KEYWORDS      HTC; cDNA; full-length; Tetraodon nigroviridis.
SOURCE        Tetraodon nigroviridis
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
              Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
              Tetraodontidae; Tetraodontidae; Tetraodon.
REFERENCE     1 (bases 1 to 1623)
AUTHORS       Genoscope.
TITLE         Direct Submission
JOURNAL       Submitted (10-AUG-2004) Genoscope - Centre National de Sequencage -
              : 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
              (E-mail : segre@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT       The sequences are based on single pass reads.
              More information available at
              http://www.genoscope.cns.fr/tetraodon.
FEATURES             Location/Qualifiers
     source           1..1623
                     /organism="Tetraodon nigroviridis"
                     /mol_type="mRNA"
                     /db_xref="taxon:99883"
                     /tissue_type="Muscle"

ORIGIN
Query Match      24.2%; Score 382; DB 3; Length 1623;
Best Local Similarity 60.7%; Pred. No. 1.1e-91;
Matches 795; Conservative 0; Mismatches 455; Indels 60; Gaps 8;

Qy 160 ATGATGACGATTTCGGCTCGCGAAATTCACATCCCGCGGTACCCAAACGTCGAGGCA 219
Db 87 ATCTGGAAGATCCACGCTCGGGAATATTCACATCCCGCGGCAACCCACCGTGGAGGTC 146
Qy 220 GAGGTTTCTCGGATGACGGTTCACCGGTGTCGAGGTGTTCCATCCGGCGCATCCACC 279
Db 147 GACCTGTACACCAAGAGGTCGTGTCAGAGCTGG---GTCCCCAGCGGCGCTCCACC 203
Qy 280 GCGCTCCACGAGGTCATGAGTCGGTGAACGGTGGCGA---TCGGTACCTGGGCAAGGCG 336
Db 204 GGCATCTACGAGGCGCTGGAGCTCGCGACAAACGACAAACCCGCTACATGGGCAAGGT 263
Qy 337 GTTTTGAAGGCAGTTGAAACGTCACGAGAAATTCGCGACGAGCTCGCTGGGCTAGAG 396
Db 264 GTCTCTAAAGCTGTTGAGAAATATCAATAAATAATTCACCTGCACTGTTGGCAAGGAC 323
Qy 397 GGTGACGATCAGCGCC-----TCATGACGAAAGCAATGATCAAGCTTTGATGGCACCGCC 450
Db 324 GTCAACGTCGTGACGACGCAAAATCGCAAGCTGATGCTGGAGATGACGCGCACCGAC 383
Qy 451 AACAAATCCCGCTCGGTCGAAACGCAATCTCTGTTGTTTCAATGGCTGTTGCAAAAGCT 510
Db 384 AACAAATCTAAGTTCCGGGCGCAACGCCATCTCTGGGCGTGTCTCTGGCTGTGTGCAAGGCT 443
Qy 511 GCTGCTGATTCGCGAGGCTCCCACTGTTCGCTACATCGGTGGACCAACAGC----- 563
Db 444 GGTGACGACGAGAGGCGGTGCCCTCTACCGCCACATCGCGACCTGGCGGCAACCCC 503
Qy 564 --ACACGTTCTTCCAGTTCCCAATGATGAACATCATCAACGGTGGCGCTCAAGTCACTCC 621
Db 504 CAAGTCATCTCTCCCGTTTCCCGCTTTCAAGGTCATCAACGGGCGGTCTCCACGACGCAAC 563
Qy 622 GGTGTTGAGTTCAAGAAATTCATGATCGTCCAATCGGTGCGAGACCTTCTCTAGGCT 681
Db 564 AAGCTGGCCATGACGAGTTATGATCTCTGCGCGGTGGGAGCGAGAGCTTCAAGAGGCC 623
Qy 682 CTCGCAACGCGCGGAGGTCTACCAAGCACTGAAAGTCCGTTCATCAAGGAAAGGCG--- 738

```

```

Db 624 ATGCGCATCGGCGCGAAGTCTACCAACCTGAAGAAGCTCATCAAGAGAGTAAGGC 683
Qy 739 ---CTGTCACCGGACTTTGGCGATGAGGCGGCTTGCTTCCTTCGCTCGGCTCCACCCGT 795
Db 684 AAGGACGCGCACCAACGTTGGGAGACGAGGAGGCTTTGCCCCCAACATCTCTGGAGAACAG 743
Qy 796 GAGGCTCTTGACCTTATCGTTGAGGCAATCGAGAAGGCTGGCTTACCCCGCAGGACAGGAC 855
Db 744 GAAGCTCTGGAGCTGCTGAAGAACGCAATCGCTTAAGCGCGGCTACAC-----CGACAAG 797
Qy 856 ATGCTCTTTCGCTTGACGCTTGCTTCTCTGAGTTCTTCAAGGACGCGCACCTTACCACTTC 915
Db 798 ATTGTGATCGGATGACGCTGCTCGAGTCTTCAAGGGCGCAAGTACGACCTG 857
Qy 916 GAA-----GGTGGCGAGCACTCCGCGAGCTCCGAGCTGAGATGGAACAGTT 954
Db 858 GACTTCAAGTCTCCGACGACCGCGCGCTACATCTCTCCGAGAGCTGGCGGACCTC 917
Qy 955 TAGCTGAGCTCGTTGACGCGTACCAATCGTCTCCATCGAGGACCCACTTGCAGGAAGAT 1014
Db 918 TACAGGGGCTTCGTCAAGATTACCCAGTGTGTCTCATCGAGGACCCCTTTGACGAGGAC 977
Qy 1015 GACTGGGAGGTTTACACCAACCTTCAACGCAACCATCGGCGACAAAGTTTCAGATCTGTTGC 1074
Db 978 GACTGGGAGGCTGGACCAATTTCAAGCCAGCACCAGC-----ATTGAGTGTGTGGC 1031
Qy 1075 GACGACTTTCGTCACCAACCTCGAGCGCTGAAGAGGCGATCGTAAAGAGGTGCC 1134
Db 1032 GAACGCTTCAAGGTCACCAACCCCAACGATCGCAAGGTTGTGCCCAAGATCTCTGC 1091
Qy 1135 AACTCCATCTGTTAAGTGAACGATCGGTACCTTCCACGAGACCTTTCGACGCTGTC 1194
Db 1092 AACTGCTGCTGCTCAAGTCAACGATCGGTTCGCTTACCGAGTCCCTGCGAGGATGC 1151
Qy 1195 GACATGGCTCACCGCGAGGCTACACTCCATCATGATGTCCACCGTTCCGCTGAGACGAG 1254
Db 1152 AAGATGGCCAGAGCAACGCTGGGCGTGTGTTGATGATCGTCCGCGGAGAGCGAG 1211
Qy 1255 GACACACCAATTGCTGACCTCGAGTTGCACTCAATGTGTGGCCAGATCAAGACTGGTGT 1314
Db 1212 GACACCTTCACTGCTGACCTCGGTGGTGTGTGTCACCGGACAGATCAAGACGCGGCGCT 1271
Qy 1315 CCAGCAGTTCCGACCGCTGTCGCAAAAGTACAAACAGCTTCTCCGCACTCGAGAGCTGCTT 1374
Db 1272 CCGTGGCTTCCGAGCGCTTGGCCAGTACAAACAGTGTCTCCGATCGAGGAAGACTG 1331
Qy 1375 GGCGACGCGCGGCTTACGCGAGGTGCGAGCGCATTTCCACGCTTTCAGGG 1424
Db 1332 GGCGACCGCGCGCTTCCGCGCAAGAACTTCAGGACCCCACTGTGAGCG 1381

```

```

RESULT 7
CR698859
LOCUS          CR698859          1613 bp      mRNA      linear      HTC 19-AUG-2004
DEFINITION    Tetraodon nigroviridis full-length cDNA.
ACCESSION     CR698859
VERSION       CR698859.1 GI:51196768
KEYWORDS      HTC; cDNA; full-length; Tetraodon nigroviridis.
SOURCE        Tetraodon nigroviridis
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
              Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
              Tetraodontidae; Tetraodontidae; Tetraodon.
REFERENCE     1 (bases 1 to 1613)
AUTHORS       Genoscope.
TITLE         Direct Submission
JOURNAL       Submitted (10-AUG-2004) Genoscope - Centre National de Sequencage -
              : 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
              (E-mail : segre@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT       The sequences are based on single pass reads.
              More information available at

```


[illegible]

```
Db 198 GGCATCTACGAGGCCCTGGAGCTCCGCGACAACGACAAAACCCGCTACATGGGCAAAAGGT 257
Qy 337 GTTTTGAAGCAGTTGAAACGTTCAACGAGAAATCGGCACAGAGCTCGCTGCGCTAGAG 396
Db 258 GTCTCTAAAGCTGTGTGAGATATCAATAAAATAATTTGSCACCTGCACCTGGTGTGGCAAGGAC 317
Qy 397 GCTGACGATCAGCGCC-----TCATCGACGAAGCAATGATCAAGCTTGTATGTCACCGCC 450
Db 318 GTCAACGTTCTGACACGAGCAAAATCGAAGCTGATGTGAGATGGACGGACCGAC 377
Qy 451 AACAAAGTCCCGCTGGGTGCAACGCAATCTTTGGTGTTCATGGCTGTGTGCAAAAGGCT 510
Db 378 AACAAATCTAAGTTTCGGGGCCACGCGCATCTGGGGGTGTCCCTGGCTGTGTGCAAGGCT 437
Qy 511 GCTGCTGATTCGGAGCGCTCCCACTGTTCGGCTAGCATCGGTGACCAAAAGC-----563
Db 438 GGTGCAGCAGAAAGGCGTGCCACTACCGGCACATCCGACCTGGCGCGCAACCCC 497
Qy 564 --ACACGTTCTCCAGTTTCCAAATGATGAACATCATCAACGTTGGCGCTCACGCTGACTCC 621
Db 498 CAAGTCACTCTCCCGTTCCGCTTTCAAGCTCATCAACGGCGGCTCCACGAGGAAAC 557
Qy 622 GGTGTGACGTTTCAGGAATTCATGATCGCTCCAATCGGTGCAGAGACCTTTCTCTGAGGCT 681
Db 558 AAGCTGGCCATGCAAGGAGTTTCATGATCTCGCGGTGGAGCGCAGCAGCTTCAAAGAGGCC 617
Qy 682 CTCGGAAACGGCGGAGGTTTACCACGCACTGAAGTCCGTCAATCAAGAAAGAGGCG---738
Db 618 ATGCGCATCGCGCGCAAGTCTACCAACCTGAAGAACGTCATCAAGGAGAAATGACGGC 677
Qy 739 ---CTGTCCACCGGACTTGGCGATGAGGGCGCTTCGCTCTCTCCGTCCGCTCCACCGCT 795
Db 678 AAGACGCCCAACACGTGGGAGAGAGGAGGCTTTGCCCCCAACATCTCTGGAGAAACAG 737
Qy 796 GAGGCTCTTGACCTTATCGTTGAGGCAATCAGAAAGCTGGCTTCAACCCAGGCAAGGAC 855
Db 738 GAAGCTCTGGAGCTGTGAAGAACGCCATCGCTAAGGCCGCGCTACAC-----CGACAAG 791
Qy 856 ATGCTCTTCTCGAGCTGTCTCTCGAGTTCTTCAGGACGCGACCTACCACTTC 915
Db 792 ATTGTGATCGGCATGGACGTGGCTGTGCTCTGAGTTCTACAAAGGCGGCAAGTACGACCTG 851
Qy 916 GAA-----GGTGGCCAGACACTCCGACGCTGAGATGGCAAAAGCTT 954
Db 852 GACTTCAAGTCTCCGAGGACCCCGCGCTACATCTCTCCGAGAGCTGGCGCACTC 911
Qy 955 TACGCTGAGCTCGTTGACGCTACCCAAATGCTTCCATTCGAGGACCCCACTGCAGGAAGAT 1014
Db 912 TACAGGGGCTTCGTCAAAGATTACCCAGTGGTGTCCATCGAGGACCCCTTTGACGAGGAC 971
Qy 1015 GACTGGGAGGTTTACCAACCTTACCGCAACCATCGCGACAGGTTTCAGATCGTTGGC 1074
Db 972 GACTGGGAGGCGTGGAACCAATTTTACAGCCAGCACCCAGC-----ATTCAAGTGGTGGGC 1025
Qy 1075 GACGACTTCTTCGTCACCAA--CCCTGAGCGCTTGAAGGAGGCGCTCGCTAAGAAAGGCTGC 1133
Db 1026 GACGACCTCAGGTCACCAACCCCAACGATCGCCAAAGGTTGGGCCAGAGTCTCTG 1085
Qy 1134 CAATCCCATCTGTTAAGGTGAACAGATCGGTACCCCTTACCCGAGACCTTCGACGCTGT 1193
Db 1086 CAATGCTGCTGTCTCAAAGTCAACAGATCGGCTCGGTTCACCGAGTCCCTTCGAGGCAATG 1145
Qy 1194 CGACATGGCTCACCGCGAGGCTACACCTCATGATGTCCACCGCTTCGGGTGAGACCGA 1253
Db 1146 CAAGATGGCCAGAGCAACCGGCTGGGGGCTGATGTTAGCCATCTGCTCCGGGAGACGGA 1205
Qy 1254 GGACACCAACCATTTGTGACCTTCGAGTTTGCACTCAACTGTGGCCAGATCAAGACTGGTGC 1313
Db 1206 GGACACCTTTCATCTGCTGACCTGGTGTGGTCTGTGCAACCGGACAGATCAAGACGGGCGC 1265
Qy 1314 TCCAGCAGTTTCCGACCGGTGTGCAAGATCAACACGCTTCTCCGCATCGAGCAGCTGCT 1373
Db 1266 TCCGTGGCGCTCCGAGCGCTTTGGGCAAGTACAACACAGCTGTCTCCGGATCGAGGAAGAGCT 1325
```

```
Qy 1374 TGCGCAGCGCGCGCTTACGAGTTCGAGGTCGAGCGCATTTCCACAGCTTTTCAGGG 1424
Db 1326 GGGCGACCAAGGCCCGTTCCCGCAAGAACTTCAGGCACCCCATCTGAGCG 1376

RESULT 9
CR728153
LOCUS CR728153 1637 bp mRNA linear HTC 19-AUG-2004
DEFINITION Tetraodon nigroviridis full-length cDNA.
ACCESSION CR728153
VERSION GI:51226414
KEYWORDS HTC; cDNA; full-length; Tetraodon nigroviridis.
SOURCE Tetraodon nigroviridis
ORGANISM Tetraodon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontoidea; Tetraodontidae; Tetraodon.
REFERENCE 1 (bases 1 to 1637)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (10-AUG-2004) Genoscope - Centre National de Sequencage -
: 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
(BE-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT The sequences are based on single pass reads.
More information available at
http://www.genoscope.cns.fr/tetraodon.
FEATURES
source
1..1637
Location/Qualifiers
/organism="Tetraodon nigroviridis"
/mol_type="mRNA"
/db_xref="taxon:99883"
/tissue_type="fish"

ORIGIN
Query Match 23.9%; Score 377; DB 3; Length 1637;
Best Local Similarity 60.4%; Pred. No. 2.5e-90;
Matches 791; Conservative 0; Mismatches 460; Indels 58; Gaps 8;

Qy 160 ATCATGACGATTTCCGTCGCGAAATTCAGACTCCCGGTAACCCACCGTCGAGGCA 219
Db 108 ATCTGAAGATCCACGCTCGGAAATATTCAGTCCCGCGCAACCCCATGGGAGGTC 167
Qy 220 GAGGTTTCTCGATGACGCTCCACGGTGTCCAGGTGTTCCATCCGCGCATCCACC 279
Db 168 GACCTGTACCAAGAAAGGTCTGTACAGAGTTCGGTCC--CCAGCGCGGCTCCACC 225
Qy 280 GGGGTCCACGAGGCTCATGAGCT--GCGTGACGGTGGCGATCGCTACTGGGCAAGGGCG 337
Db 226 GGCATCTACGAGGCCCTGGAGCTCCGCAACGACAAACCCGCTACATGGGCAAGGTG 285
Qy 338 TTTTGAAGGAGTTGAAAACGTCAAAGAAATTCGCGACGAGCTCGTGGCTAGAGG 397
Db 286 TCTCTAAGCTGTGTGAATATCAATAAAATAATTGCACCTGCTGGTGTGGCAAGGACG 345
Qy 398 CTGACGATCAGCGCC-----TCATCGACGAAGCAATGATCAAGCTTGTAGTGGACCCGCA 451
Db 346 TCAACGTTCTGGACAGGCAAAATCGAAGCTGATGCTGGAGATGCAACGACCCGACA 405
Qy 452 ACAAGTCCCGCTCGGTGCAAAAGCAATCTTGGTGTTCATGGCTGTGTGCAAGGCTG 511
Db 406 ACAATCTAAGTACCGGGGCAACGCCATCTCGGGCGTGTCTGGCTGTGTGCAAGGCTG 465
Qy 512 CTGCTGATTCCGAGGCTCCCACTGTTCGCTATCATCGGTGGACCAAGC-----563
Db 466 GTGCAGCAGAAAGGCGTGGCCCTCTACCGCCACATCGCCGACCTGGCGGCAACCCCC 525
Qy 564 -ACACGTTCTTCCAGTTTCCAAATGATGAACATCATCAACGGTGGCGCTCACGCTGACTCG 622
Db 526 AAGTCATCTCCCGTTTCCCGCTTTTCAACGTCATCAACGCGGCTCCACGACGGAACA 585
Qy 623 GTGTTGAGTTTCAGGAATTCATGATCGCTCAATCGGTGCAGAGACCTTCTCTGAGGCTC 682
```

```
Db 586 AGCTGGCCATCAGGAGTTCATGATCTCTCCCGGTGGAGCCAGCAGCTTCAAGAGGCCA 645
Qy 683 TCCGCAACGGCGCGAGGTCTTACACGCACTGAAGTCCGTCAATCAAGGAAAGGCG--- 738
Db 646 TGCATCGGCGCGAGTCTACCAACCTGAAGAGCTCATCAAGAGAGTACGGCA 705
Qy 739 --CTGTCCACCGGACTTGGCGATGAGGCGGCTTTCGTCTCTTCGTCCGTCCACCCGCTG 796
Db 706 AGGACGCCACCAACGTTGGGAGACGAGGAGGCTTTGCCCCCAACATCTGGAGAACAGG 765
Qy 797 AGGCTCTTGACCTTATCGTTGAGGCAATCGAAGGCTGGCTTCAACCCAGGCAAGGACA 856
Db 766 AAGCTCTGGAGCTCTGAAGAACGCCATCGCTAAGGCGCGGTACAC-----CGACAAGA 819
Qy 857 TCGCTCTTGCTCTGAGCTTCTCTCTGAGTCTTCAAGGAGCGGACCTTACACATTCG 916
Db 820 TTGTGATCGGATGAGCTGCTGCTCCGAGTCTTCAAGGCGGCAAGTACGACCTGG 879
Qy 917 AA-----GGTGGCCAGCACTCCGCACTGAGATGCAACAGTTT 955
Db 880 ACTTCAAGTCTCCCGAGACCCCGCGCTACATTTCTCCGAGCAGCTGSCCGACCTCT 939
Qy 956 ACGCTGAGCTGTTGACGCGTACCAATCGTCTCTCATCGAGGACCCACTGAGGAAGATG 1015
Db 940 ACAGGGGCTTCTCAAGATTAACCCAGTGTGTCTCATCGAGGACCCCTTTGACAGGACG 999
Qy 1016 ACTGGGAGGTTACACCAACCTCAGCGCAACCATCGGCGACAAAGTTTCAGATCTTGGCG 1075
Db 1000 ACTGGGAGGCGTGACCAATTTTACGCCAGCACGAGC-----ATTGAGTGTGGCG 1053
Qy 1076 ACAGCTTCTTCTTCCACCAACCTCAGCGCTGGAAGGAGGCGATCGCTAAGAGGCTGCA 1135
Db 1054 ACGACCTTACGCTCACCAACCCCAACGCAATCGCCAGGCTGTGCCAGAGTCTCTGCA 1113
Qy 1136 ACTCATCTCTGTTAAGTGAACAGATCGGTACCTCAACCGAGACCTTCGACGCTGTG 1195
Db 1114 ACTGCTCTGCTCTCAAGTCAACAGATCGGCTCGGTCAACGAGTCTCTCGAGCATGCA 1173
Qy 1196 ACATGCTCACCGCGAGGCTACACCTCCATGATGTCCACCGTTCGGTGAGACCGAGG 1255
Db 1174 AGATGGGCCAGAGCAACGGCTGGGCGGTGATGGTCAGCCTATCGCTCCGGGAGACGAGG 1233
Qy 1256 ACACCACTTGTGACCTCGCAGTTCGACCTCAACTGTGGCCAGATCAAGACTGTGCTC 1315
Db 1234 ACACCTTATCGCTGACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1293
Qy 1316 CAGCAGCTTCGACCGGTGTGCAAGTACAAACAGCTTCTCCGATCGAGCAGCTGCTTG 1375
Db 1294 CGTCCGCTCGAGCGCTTGGCCAAAGTACAAACAGCTGTCTCCGATCGAGGAGCTGG 1353
Qy 1376 GCGAGCGGGGCTTACGAGGTGCGAGCGATTCACCGATTCACCGTTTCAGGG 1424
Db 1354 GCGACGAGCGCGCTTCCGCGCAAGAACTTCAGGACCCCATCTGAGCG 1402

RESULT 10
CR685932 LOCUS 1551 bp mRNA linear HTC 19-AUG-2004
DEFINITION Tetraodon nigroviridis full-length cDNA.
ACCESSION CR685932
VERSION CR685932.1 GI:51183839
KEYWORDS HTC; cDNA; full-length; Tetraodon nigroviridis.
SOURCE Tetraodon nigroviridis
ORGANISM Tetraodon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodontidae; Tetraodon.
REFERENCE 1 (bases 1 to 1551)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (10-AUG-2004) Genoscope - Centre National de Sequencage
```

```
: 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
(E-mail : seque@genoscope.cns.fr - Web : www.genoscope.cns.fr)
The sequences are based on single pass reads.
More information available at
http://www.genoscope.cns.fr/tetraodon.
Location/Qualifiers
source 1.1551
/organism="Tetraodon nigroviridis"
/mol_type="mRNA"
/db_xref="taxon:99883"
/tissue_type="Eyes"

Query Match 23.4%; Score 369.6; DB 3; Length 1551;
Best Local Similarity 60.3%; Pred. No. 2.4e-88;
Matches 782; Conservative 0; Mismatches 454; Indels 60; Gaps 8;

Qy 174 CGCTCGGGAATTTTCGACTCCCGCGTAAACCCAGCTGAGGAGAGGTTTCTTGA 233
Db 27 CGCTCGGGAATTTTCGACTCCCGCGCAACCCACCGTGGAGTTCGACCTGTACACCA 86
Qy 234 TGAGGTTTCCACGCTGTCGAGGTTCATCCGCGCATCCACCGCGTCCACGAGGC 293
Db 87 GAAAGGTCTGTTTCAAGAGCTGCG---GTCCCGAGCGCGCTCCACCGCATCTACGAGGC 143
Qy 294 TCATGAGCTCGCTG---ACGCTGGCGATCGCTACTGGGCAAGGCGCTTTTGAAGGCGAGT 350
Db 144 CTTGAGCTCCGCGACAAACGATTAAACCCGCTACATGGCAAAAGGGTCAAAAGGCGCT 203
Qy 351 TGAAGCGTCAACGAAGAAATTCGCGAGCAGCTCGCTGGGCTAGAGGCTGACGATCAGCG 410
Db 204 TAAATATATCAATGAGTTTGGCCCGCTTGTGTAACCGAGCAGCTCAACGCTTCTGGA 263
Qy 411 CC-----TCATCGACCAAGCAATGATCAAGCTTGATGGCACCCGCCAACAAGTCCCGCT 464
Db 264 CCAGGCAAAATCGACAAGCTGATGCTGGAGATGAGCGCACCCGCAACAACAACTTAAGTT 323
Qy 465 GGGTGCAAAACGCAATCTTGGTGTTCATGGCTGTTGCAAAAGGCTGCTGCTGATTCCGC 524
Db 324 CGGGGCCAACGCCATCTGGCGGTGCTTCCGTGGCTGTGTGCAAGGCTGTGTCAGAGAGAA 383
Qy 525 AGGCCTCCCACTGTTCGCTACATCGGTGGACCAACAG-----CACACGTTCTTTC 575
Db 384 GGGGTGCTCCCTCTACCGCCACATTTGCGACCTTGGCGGCAACCCCAAGTCATCTCTCC 443
Qy 576 AGTTCGAATGATGAACATCATCAACGTTGGGCTCAGCTGACTCCGGTGTGAGTTCA 635
Db 444 CGTTCCGCTTTCAACGCTGATCAACGCGGCTCCACGCGAGCAACAGCTGGCCATGCA 503
Qy 636 GGAATTCATGATCGCTCCAATCGTGCAGAGACCTTCTCTGAGGCTCTCCGCAACGGCGC 695
Db 504 GAGTTTATGATCTTCCCGTGGAGCCAGAGCTTCAAGAGGCGCATTCGCATTCGGGCG 563
Qy 696 GAGGCTTACACGCACTGAAGTCCGCTCATCAAGAAAAGGCG-----CTGTCCACCGG 749
Db 564 CGAGGCTACCAACAACCTGAAGACGTCATCAAGGAGAGTACGCGCAAGGACGCCACCA 623
Qy 750 ACTTGGCGATGAGCGCGCTTTCGCTCTTCGTCGGCTCCACCGCTGAGGCTCTTGACCT 809
Db 624 CGTGGAGAGACGAGGAGGCTTTGCCCCCAACATCTCTGGAGAACAAAGGAAGCTCTGGAGCT 683
Qy 810 TATCGTTGAGGCAATCGAAGGCTGGCTTTCACCCAGGCAAGACATCGCTCTTGCTCT 869
Db 684 GCTGAAGAACGCCATCTGCTAAGGCGGCTACAC-----CGACAAGATTGTGTCGGCAT 737
Qy 870 GAGCGTTGCTTCTCTGAGTTCTTCAAGAGCGGCACTTACACCTTCGAA----- 918
Db 738 GAGCGTGGCTTCCGCTTCTACAGGGGCGCAAGTACGACTGACTTCAAGTCCCC 797
Qy 919 -----GGTGGCCAGCACTCCGCGATGAGATGGCAAAAGTTTACGCTGAGCTCGT 968
Db 798 CGACGACCCCGCGCTTACATCTCTCCGAGCAGCTGGCCGACCTCTACAGGGGCTTCGT 857
```


Qy 1313 CTCGACGACGTTCCGACCGTTCGCAAGTACACAGCTTCTCCGATCGAGCGTGC 1372
 |||||
 Db 1293 CTCGTCGCGCTCCGAGCGCTGGCCAGTACACAGCTCTCCGATCGAGGAGGC 1352
 |||||
 Qy 1373 TTGGCGACCGCGCGTCTACCGAGTGCACGCGATTCCTCCAGCTTTTCAGGG 1424
 |||||
 Db 1353 TGGCGGACCGAGCGCCCTTCGCGGCAAGAACTTCAGGCACCCATCTGAGCG 1404
 |||||

RESULT 12

CR656482 1582 bp mRNA linear HTC 18-AUG-2004
 LOCUS
 DEFINITION Tetraodon nigroviridis full-length cDNA.

CR656482

VERSION CR656482.1 GI:51152927

KEYWORDS HTC; cDNA; full-length; Tetraodon nigroviridis.

SOURCE Tetraodon nigroviridis

ORGANISM Tetraodon nigroviridis

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 Acanthomorpha; Acanthopterygii; Percormorpha; Tetraodontiformes;
 Tetraodontidae; Tetraodontidae; Tetraodon.
 1 (bases 1 to 1582)

REFERENCE

AUTHORS Genoscope.

TITLE Direct Submission

JOURNAL Submitted (10-AUG-2004) Genoscope - Centre National de Sequencage -
 : 2 rue Gaston Cremlieux, CP 5706 - 91057 EVRY cedex - FRANCE
 (E-mail : seque@genoscope.cns.fr - Web : www.genoscope.cns.fr)

COMMENT

The sequences are based on single pass reads.

More information available at

http://www.genoscope.cns.fr/tetraodon.

FEATURES

Location/Qualifiers

1..1582

/organism="Tetraodon nigroviridis"

/mol_type="mRNA"

/db_xref="taxon:99883"

/tissue_type="Liver"

ORIGIN

Query Match 23.1%; Score 364.6; DB 3; Length 1582;
 Best Local Similarity 60.3%; Pred. No. 5.5e-87;
 Matches 791; Conservative 0; Mismatches 459; Indels 61; Gaps 9;
 Qy 160 ATCATGACGATTTTCGCTCGGAAATTCGACCTCCGCGTAAACCAACCGTCGAGCA 219
 |||||
 Db 67 ATCTGAGATCCAGCTCGGAAATTTGACATCCCGCGCAACCCACCGTGGAGTC 126
 |||||
 Qy 220 GAGTTTTTCCTGGATACGGTTCCACGCGTTCGACGGTGTTCATCCGCGCATCCACC 279
 |||||
 Db 127 GACCTGTACACCAAGAAAGTCTGTTACAGAGTGGC---GTCCCGAGCGCGCTCCACC 183
 |||||
 Qy 280 GCGTTCACGAGGCTCATGCTCGGTGACGGTGGCA---TCGCTACCTGGGCAAGGC 336
 |||||
 Db 184 GGCATCTACGAGGCGCTGGAGCTCCGCGAACACACAAACCCGCTACATGGGCAAGGT 243
 |||||
 Qy 337 GTTTTGAAGCGAGTTGAAACGCTCAACGAAGAAATCGCGACGAGCTCGCTGGCTAGAG 396
 |||||
 Db 244 GTCTTAAGCTGTGTGAGATATCAATAAATAATTTGACCTGCACTGGTTGGCAAGGAC 303
 |||||
 Qy 397 GCTGACGATCAGCGCC-----TCATCGACGAAGCAATGATCAAGCTTGTATGGCACCGCC 450
 |||||
 Db 304 GTCAACGTTCTGGACAGGCAAAATTCGACAAAGCTGATCTGGAGATGGACGGCACCGAC 363
 |||||
 Qy 451 AACAAAGTCCCGCTGGGTGCAACGCAATCTTGGTGTTCATGGCTGTTCGCAAGGCT 510
 |||||
 Db 364 AACAAATCTAAACACGGGCGCAACGCCATCTCGGCGTGTCCCTGGCTGTGTGCAAGGCT 423
 |||||
 Qy 511 GCTGCTATTCCGCGAGCGCTCCCACTGTTCGCTACATCGTGGACCAAGC----- 563
 |||||
 Db 424 GGTGCACAGAAAGGCGGTGCGCACTTACCGCCACATCGCCGACCTGGCGGCAACCCC 483
 |||||
 Qy 564 --ACACGTTCTTCAGTTCCAATGATGAACATCATCAACGGTGGCGCTCAGCTGACTCC 621
 |||||

Db 484 CAAGTCATCTCCCGCTTCGCGTTTCAAGTCAATCAACGCGCGCTCCACGCAAGAAC 543
 |||||
 Qy 622 GGTGTTGACGTTGAGGAATTCATGATCGTCCCAATCGGTGAGAGCTTCTCTGAGGCT 681
 |||||
 Db 544 AAGTGGCCATGACGAGGATTCATGATCTCGCGGTGGAGCCAGCAGCTTTCAAAGAGCC 603
 |||||
 Qy 682 CTCGCAACCGCGCGGAGGCTTACACGACCTGAAGTCCGTCATCAAGAGAAAGGCG--- 738
 |||||
 Db 604 ATGCGCATCGCGCGGAGTCTACCAAACTGAAGAACTCATCAAGGAGAGTACGGC 663
 |||||
 Qy 739 ---CTGTCCACCGGACTTTGGGATGAGGGCGCTTCCTCTCCGTCGGCTCCACCGT 795
 |||||
 Db 664 AAGGACGCCCAACGCTGGGAGACGAGGAGGCTTTGCCCAACATCTCTGGAGAACAG 723
 |||||
 Qy 796 GAGGCTCTTACCTTATCGTTGAGGCAATCGAAGGCTGGCTTACCCAGGCAAGGAC 855
 |||||
 Db 724 GAAGCCCTGGAGCTGCTGAAGAACGCCATCGTAAAGCCGCTACAC-----CGACAAG 777
 |||||
 Qy 856 ATCGCTCTGCTCGGACGTTGCTTCTCTGAGTTCTTCAAGGACGGCACCTACACATTC 915
 |||||
 Db 778 ATTGTGATCGCATGGAGTGGCTGCTCTGAGTTCTACAAGGGCGCAAGTACGACCTG 837
 |||||
 Qy 916 GAA-----GGTGGCAGCACTCCGCGAGCTCCGAGCTGAGATGGCAAAAGTT 954
 |||||
 Db 838 GACTTCAAGTCTCCCGACGACCCCGCGCTACATCTCTCCGAGAGCTGGCGACCTC 897
 |||||
 Qy 955 TAGCTGAGCTCGTTGACGCTACCAATCGTCTCAATCGAGGACCCACTTCGAGGAGAT 1014
 |||||
 Db 898 TACAGGGGCTTCGTTCAAAAGATTACCCAGTGGTTCATCGAGGACCCCTTTTGACAGGAC 957
 |||||
 Qy 1015 GACTGGGAGGTTTACACCAACCTCACGCAACCATCGCGACAAAGTTTCAGATCGTTGGC 1074
 |||||
 Db 958 GACTGGGAGGCTGGACCAATTTCAAGCCAGCAGCAGC-----ATTGAGTGGTGGC 1011
 |||||
 Qy 1075 GAGGACTTCTTCGTCACCA--CCTGAGCGCTGAAGAGGAGCATCGCTAAGAGGCTGC 1133
 |||||
 Db 1012 GAACGCTCAGCGTCAACCAACCCCAACGATCGCAAGGGTGTGGCCCAAGAGTCTCTG 1071
 |||||
 Qy 1134 CAATCTCATCTGTTAAGGTGACGATCGGTACCTCCTCACGAGACCTTCGACGCTGT 1193
 |||||
 Db 1072 CAATGCTCTGCTCAAGATCAACAGATCGGCTCGGTCAACGAGTCCCTGCGAGGATG 1131
 |||||
 Qy 1194 CGACATGGCTCACCGCGAGCTACACCTCATGATGTCCACCGTTCGCGGTGAGACCGA 1253
 |||||
 Db 1132 CAAGATGGCCGAGGCAACCGCTGGGCGGTGATGGTCAGCATCGCTCCGGGAGAGCGA 1191
 |||||
 Qy 1254 GGACACCACTTCTGACCTCGAGTTGCACTCAACTGTGGCCAGATCAAGACTGTGTGC 1313
 |||||
 Db 1192 GGACACCTTCTGTTCTCTCTGTTGGTGGTCTGTGCAACCGACAGATCAAGACGGGCGC 1251
 |||||
 Qy 1314 TCCAGCAGCTTCGACCGTGTGCAAGATCAACAGCTTCTCCGATCGAGCAGCTGCT 1373
 |||||
 Db 1252 TCCGTGCGCTCCGAGCGTTGGCAAGTAAACAGCTGCTCCGGATCGAGGAGAGCT 1311
 |||||
 Qy 1374 TGGCGACGCGCGGTCTTACGAGCTCGCAGCGATTCGCCAGCTTCCAGGG 1424
 |||||
 Db 1312 GGGCGACGAGCGCGTTCGCGGCAAGAACTTCAGGCACCCCATCTGAGCG 1362
 |||||

RESULT 13

BH770828

LOCUS

DEFINITION

LMGag569 MG1363 Random Sequence Tag Library Lactococcus lactis

subsp. cremoris genomic, genomic survey sequence.

BH770828

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

BH770828 2021 bp DNA linear GSS 01-MAY-2002
 LMGag569 MG1363 Random Sequence Tag Library Lactococcus lactis
 subsp. cremoris genomic, genomic survey sequence.

BH770828

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

Lactococcus lactis subsp. cremoris
 Lactococcus lactis subsp. cremoris
 Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 Lactococcus.
 1 (bases 1 to 2021)

AUTHORS Bolotin, A., Ehrlich, S. D. and Sorokin, A.
TITLE Studies of genomes of dairy bacteria *Lactococcus lactis*
JOURNAL Sci. Aliments (2002) in press
COMMENT Contact: Sorokin A
 Genetique Microbienne
 INRA
 CRJ INRA, Domaine de Vilvert, 78352 Jouy en Josas cedex, France
 Tel: 33 1 34 65 25 16
 Fax: 33 1 34 65 25 21
 Email: sorokine@jouy.inra.fr
 best homologue in strain IL1403 is enoB (95%)
 Class: shotgun
 High quality sequence start: 30
 High quality sequence stop: 1993.
FEATURES Location/Qualifiers
 source
 1. .2021
 /organism="Lactococcus lactis subsp. cremoris"
 /mol_type="genomic DNA"
 /strain="MGL1363"
 /sub_species="cremoris"
 /db_xref="taxon:1359"
 /clone_lib="MGL1363 Random Sequence Tag Library"
 /note="vector: pSGM2; Site 1: SmaI; Library of chromosomal fragments of L.lactis strain MGL1363 was prepared by partial AluI digestion or by sonication."
ORIGIN
 Query Match 21.7%; Score 341.8; DB 8; Length 2021;
 Best Local Similarity 56.9%; Pred. No. 8.4e-81;
 Matches 691; Conservative 0; Mismatches 512; Indels 12; Gaps 3;
 Qy 159 AATCATGACGTATTCGTCGGAATTCCTCGACTCCGCGGTAAACCCACCGTCGAGGC 218
 Db |||||
 Qy 709 AATCGAAACATTCACGCAAGAGAAATTTTGGATTCCAGAGGAAATCCAAACCGTTGAGGT 768
 Db |||||
 Qy 219 AGAGGTTTCTCGTGAACGGTTCCACGGTGTCCAGAGTGTTCATCCGCGCATCCAC 278
 Db |||||
 Qy 769 AGATGTTAGATTGACAGATGAACTTTGGGACGGCGGCAGTTCCTCAGAGGATCAAC 828
 Db |||||
 Qy 279 CGCGCTCCACGAGCTCATGAGTCGCGTGCAGGTGGCGATCGTACCTGGGCAAGGCGGT 338
 Db |||||
 Qy 829 AGGAGATCGTGAAGCGGTGGAATTAAGAGATGGCGGAGCGGACTTCAAGGAAAGCGGT 888
 Db |||||
 Qy 339 TTTGAAGCGAGTTGAACGTTCAACGAGAAATCGGCGACGAGCTCGCTGCGCTAGAGGC 398
 Db |||||
 Qy 889 TTCTAAAGCTGTGCAATGTTAATGGTGAAATTTATGAAGCCTTTAAAGGCCAATCACC 948
 Db |||||
 Qy 399 TGAGCATCAGCGCTCATCGACGAGCAATGATCAAGCTTTCATGGCACCGCCAAACAGTC 458
 Db |||||
 Qy 949 ATTCAATCAAGCAAAATTAGACCATCTCATGATTTGAACCTTGATGGGACAAAATAAGTC 1008
 Db |||||
 Qy 459 CCGCTGGGTGCAACGCAATCTTGGTGTTCATGGCTGTTGCAAAAGGTGCTGTGA 518
 Db |||||
 Qy 1009 TCGCTTAGTGGCAATGCAATATTGGGTGTTTCCATGGCTATTGCTGCGCGCAGCTAA 1068
 Db |||||
 Qy 519 TTCGCGAGCTCCACGTGTTCCGCTACATCGGTGGACCAACGACACGCTTCTCCAGT 578
 Db |||||
 Qy 1069 CAGTGAAGAAATCCACATTTATCGTTATCTTGGTGGAGTAGATTGGAACTTCCTCAA -- 1126
 Db |||||
 Qy 579 TCCAATGATGAACATCATCAACGGTGGCGCTCAGCGTACCTCCGCTGTTGACGTTTCAGGA 638
 Db |||||
 Qy 1127 -CCATTTTCAATGTCATCAATGGTGGAGTAGCAGCCGATTCAGGAATTGACGTTCAAGA 1185
 Db |||||
 Qy 639 ATTCAATGATCGCTCAATCGGTGACAGACCTTCTCAGAGGCTTCCGCAACGGCGCGGA 698
 Db |||||
 Qy 1186 ATTCTTGATTACCCCTGTCAAAACGTGCTAGTGTTCGCGATGGTTTGAAGAAATTCGCAA 1245
 Db |||||
 Qy 699 GGTCTACACGCACTGAAGTCGTCATCAAGGAAAGGCGCTGTCCACCGGACTTGGCGA 758
 Db |||||
 Qy 1246 TATCTATCATCTTTGAAGAAATTTTATAGCTGATTAAGACTAGAAACGGCGGTAGGGGA 1305
 Db |||||
 Qy 759 TGAGGGCGGCTTCGCTCCTTCGTCGCGCTCCACCGGTGAGGCTCTTGACCTTATGCTTGA 818
 Db |||||

Db 1306 CGAGGGTGCTTTTGTCTCCAAACTTGGTTCGACGGAATAATGCAATGTCTACACTTTATCA 1365
 Qy 819 GGCATTCGAGAAGGCTGGCTTCACCCAGGCAAGACATCGCTCTTGTCTCTGACGTTGC 878
 Db |||||
 Db 1366 AGCAATTGAAAGAGCTGGTTATGTGCCAGCGGAAGAAATTCGGAATTCGATTCACCCGGC 1425
 Qy 879 TTCCTCTGAGTTCTT-----CAAGGACGGCACCTTACCACCTTCGAAGGTGGCCGACACTC 932
 Db |||||
 Db 1426 TTCAGCGAATTTATGATGACAAAGAAAGTCTATCATTTTGAAGGTCAAAAGCTCAC 1485
 Qy 933 CGCAGCTGAGATGGCAAAAGCTTTACGCTGAGCTGTTGACGGGTACCCAAATCG---TCTC 989
 Db |||||
 Db 1486 TTCAGCAGAAATGTTGACCTAAATATGAGGGTTAGTTGAAAAATATCTGCCCTAATCTC 1545
 Qy 990 CATCGAGGACCCACTGCAGGAGATGACTGGGAGGTTACACCAACTCACCGCAACCAT 1049
 Db |||||
 Db 1546 TATTGAAGACGGATTTTCGGAACATGATTGGGCTGGTTTTCGGGCTCAACGAAAGTTCA 1605
 Qy 1050 CGCGCAAGAGTTTCAGATCGTTGGCGACGACTTCTTCGTCCACCAACCTTGAGCGCTGAA 1109
 Db |||||
 Db 1606 AGTCAAAATAATTCATTTGGTGTATGATATTTTGTACAAATCCCGAAATCTTCAA 1665
 Qy 1110 GGAGGCAATCGCTAAGAGGCTGCCAATCCATCTCTGTTAAGGTGAACAGATCGGTAC 1169
 Db |||||
 Db 1666 AGAAGGAATTCAAAAAAGTGTAGCAAAATGCTATTTTGAATTAATACTAAATCAATCGGAC 1725
 Qy 1170 CCTACCGGAGACCTTCGACGCTGTCACATGGCTCACCGGCGAGGCTACACCTCCCATGAT 1229
 Db |||||
 Db 1726 AGTGACCGAAGCAATTTGAAGCGATTAGTTTGGCCAGAAAAGAGGTTTATAAGACCATGAT 1785
 Qy 1230 GTCCCAACCGTTCCCGGTGAGACCGGAGACACCACTTGTCTGACCTCGCAGTTGCACCTCAA 1289
 Db |||||
 Db 1786 TTCTCATCGTTAGAGTGAAACGGTTGATAGTTTATTTGCTGACTTTGCTATAGCCATGCA 1845
 Qy 1290 CTGTGCCAGATCAAGACTGGTCTCCAGACGCTCCGACCGTGTGCGAAAAGTACAACCA 1349
 Db |||||
 Db 1846 CGCTGCAAAATAAAAAACAGGTTCAATGGCGAAGTGACGGGTTGAAAAATATAATCA 1905
 Qy 1350 GCTTCTCCGATCGA 1364
 Db |||||
 Db 1906 ATTTTGGCGATTGA 1920
RESULT 14
CR658829
LOCUS Tetraodon nigroviridis full-length cdna. linear HTC 18-AUG-2004
DEFINITION Tetraodon nigroviridis full-length cdna.
ACCESSION CR658829
VERSION CR658829.1 GI:51155274
KEYWORDS HTC; cDNA; full-length; Tetraodon nigroviridis.
SOURCE Tetraodon nigroviridis
ORGANISM Tetraodon nigroviridis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
 Tetraodontoidea; Tetraodontidae; Tetraodon.
REFERENCE 1 (bases 1 to 1551)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (10-AUG-2004) Genoscope - Centre National de Sequencage -
 : 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
 (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT The sequences are based on single pass reads.
 More information available at
 http://www.genoscope.cns.fr/tetraodon.
FEATURES Location/Qualifiers
 source
 1. .1551
 /organism="Tetraodon nigroviridis"
 /mol_type="mRNA"
 /db_xref="taxon:99883"
 /tissue_type="Muscle"

Db	1098	CTGCTGCTCAAGGTCAAC	CGAGTCCGCTCAGT	GACCGAGTCCATCA	AAAGCGTGTAA	CGTGTAAAGCTG	1157
Qy	1201	GCTCACCGCGCAGGCTAC	ACTCATGATGTC	CCACCGTTC	CGGTGAGACCG	GAGACACC	1260
Db	1158	GCCAGAGCAGTGGT	TGGGGTGTATGGT	CAGCCATCGCT	CTGGAGAACTG	GAGACACC	1217
Qy	1261	ACCATTGCTGACCTCG	CAGTTGCAC	TCACTCACT	TGTGGCCAGATCA	AGACTGTTGCTCCAGCA	1320
Db	1218	TTCATCTCCGACCTGT	TGGTGTGACT	TGACCGGACAGAT	CAAGACGGCG	CGCTCCCTGC	1277
Qy	1321	CGTTCGACCGTGT	CGAAAGTAC	ACCAGCTTCT	CCGCATCGCAGCAGCT	GCTTTGGCGAC	1380
Db	1278	AGGTCGGAGCGTT	TGGCCAAGTACA	CCAGCTGATG	AGATCGAGGAA	GAGCTTTGGGAC	1337
Qy	1381	GCCGGCGCTAC	GCGAGT	1398			
Db	1338	AAGCCAA	GTTCGCGGT	1355			
RESULT 15							
CR733357							
LOCUS	CR733357	1407 bp	mRNA	linear	HTC	19-AUG-2004	
DEFINITION	Tetraodon nigroviridis full-length cDNA.						
ACCESSION	CR733357						
VERSION	CR733357.1 GI:51231674						
KEYWORDS	HTC; cDNA; full-length; Tetraodon nigroviridis.						
SOURCE	Tetraodon nigroviridis						
ORGANISM	Tetraodon nigroviridis						
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes; Tetraodontoidea; Tetraodontidae; Tetraodon.						
REFERENCE	1. (bases 1 to 1407)						
AUTHORS	Genoscope.						
TITLE	Direct Submission						
JOURNAL	Submitted (10-AUG-2004) Genoscope - Centre National de Sequencage - 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)						
COMMENT	The sequences are based on single pass reads. More information available at http://www.genoscope.cns.fr/tetraodon.						
FEATURES	source						
	1. 1407						
	Location/Qualifiers						
	/organism="Tetraodon nigroviridis"						
	/mol_type="mRNA"						
	/db_xref="taxon:99883"						
	/tissue_type="fish"						
ORIGIN							
	Query Match 20.7%; Score 327; DB 3; Length 1407;						
	Best Local Similarity 59.7%; Pred. No. 7.6e-77;						
	Matches 686; Conservative 0; Mismatches 410; Indels 53; Gaps 6;						
Qy	317	ATCGCTACTCTGGGCAAGGCGGCTTT	TGAAGCAGT	TGAAAA	CGTCAACGA	-----AGAAAT	371
Db	13	ATGCGTACTGGGCAAAAGGGGT	CAAAAGSCCGTTAAATATGT	CAATGAGT	TTTTTGCCC	72	
Qy	372	CGCGGAGGAGCTCGCTGGCCCTAG	AGGCTGACGATCAGCGCCTCAT	CGACGAAGCAATGAT	431		
Db	73	CGCGCTTGTGTAAACGACGAGC	GTCAACGCTTCGACCCAGGCAAAAT	TCGCAAGCTGATGCT	132		
Qy	432	CAAGCTTGTATGGCACCGCCAA	CAAGTCCCGCTGGGTGCAAA	CGCAATCCTTGGTGTTC	491		
Db	133	GGAGATGGACGGCACCGAACAA	CAATTAAGTTGGGGCCCAACGCCAT	CTCTGGCGTGTCT	192		
Qy	492	CATGGCTGTTTGC	AAAGGCTCTGCTGAT	TCCGAGGCTCCCACTGTTCCGCTACATCGG	551		
Db	193	CCTGGCTGTGTGAAGGCTG	TGTCGACAGAGAAGGCGTGC	CCCTCTACCCGCACATCGC	252		
Qy	552	TGGACCAACG	-----CACAGCTTCTTC	CAGTTTCCAATGATGAACATCATCAACGG	602		
Db	253	CGACCTGGCGGCAACCCCAAG	TCATCTCTCCCGCTTCAACGCTCATCAACGG	312			

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 29, 2005, 09:39:18 ; Search time 38.5 Seconds
(without alignments)
824.048 Million cell updates/sec

Title: US-10-728-947-2
Perfect score: 2155
Sequence: 1 VAEIMHVFAREILDSRGNT.....QLLDAGVYAGSAPFRFQG 425

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*

- 1: /cgn2_6/ptodata/1/iaa/5A-COMB.pep.*
- 2: /cgn2_6/ptodata/1/iaa/5B-COMB.pep.*
- 3: /cgn2_6/ptodata/1/iaa/6A-COMB.pep.*
- 4: /cgn2_6/ptodata/1/iaa/6B-COMB.pep.*
- 5: /cgn2_6/ptodata/1/iaa/6C-COMB.pep.*
- 6: /cgn2_6/ptodata/1/iaa/6D-COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2155	100.0	425	4	US-09-860-768-2
2	2151	99.8	425	4	US-09-860-768-4
3	1361.5	63.2	432	4	US-09-902-540-12060
4	1302.5	60.4	434	4	US-09-710-279-2322
5	1302.5	60.4	440	3	US-09-134-001C-5641
6	1282.5	59.5	444	4	US-09-252-991A-21020
7	1281.5	59.5	434	4	US-09-830-217-22
8	1281.5	59.5	434	4	US-10-278-946-22
9	1256.5	58.3	434	4	US-09-583-110-3817
10	1244	57.7	398	4	US-09-107-532A-3810
11	1232.5	57.2	444	4	US-09-328-352-7269
12	1226.5	56.9	420	3	US-08-847-065-19
13	1226.5	56.9	420	4	US-09-829-382-19
14	1226	56.9	459	4	US-09-489-039A-10061
15	1202.5	55.8	445	4	US-09-540-236-2354
16	1190.5	55.2	456	4	US-09-543-681A-6090
17	1167.5	54.2	434	4	US-09-438-185A-802
18	1131.5	52.5	407	4	US-09-107-433-4986
19	1101	51.1	443	4	US-09-949-016-8359
20	1099	51.0	434	4	US-09-949-016-6153
21	941.5	43.7	452	4	US-09-198-452A-853
22	902	41.9	311	3	US-08-961-083-184
23	902	41.9	311	4	US-09-536-784-184
24	654.5	30.4	301	4	US-08-248-796A-17053
25	504.5	23.4	177	1	US-08-446-920-11
26	458	21.3	148	4	US-09-248-796A-17054
27	265	12.3	113	4	US-09-513-999C-5420

28	224	10.4	84	4	US-09-198-452A-854	Sequence 854, Ap
29	214.5	10.0	70	3	US-08-847-065-17	Sequence 17, Appl
30	214.5	10.0	70	4	US-09-829-382-17	Sequence 17, Appl
31	214	9.9	80	4	US-09-513-999C-5931	Sequence 5931, Ap
32	206	9.6	74	4	US-09-513-999C-5841	Sequence 5841, Ap
33	181.5	8.4	70	3	US-08-847-065-16	Sequence 16, Appl
34	181.5	8.4	70	4	US-08-829-382-16	Sequence 16, Appl
35	156	7.2	48	3	US-08-715-034-2	Sequence 2, Appli
36	146	6.8	706	4	US-09-252-991A-30575	Sequence 30575, A
37	146	6.8	956	4	US-09-252-991A-17124	Sequence 17124, A
38	143	6.6	50	3	US-08-715-034-1	Sequence 1, Appli
39	119.5	5.5	718	4	US-09-252-991A-31865	Sequence 31865, A
40	119.5	5.5	1569	4	US-09-711-164-312	Sequence 312, App
41	117.5	5.5	523	4	US-09-252-991A-17991	Sequence 17991, A
42	112	5.2	472	4	US-09-134-000C-3893	Sequence 3893, Ap
43	112	5.2	1174	4	US-09-252-991A-29279	Sequence 29279, A
44	111.5	5.2	579	4	US-09-252-991A-29372	Sequence 29372, A
45	111	5.2	295	4	US-09-134-000C-4352	Sequence 4352, Ap

ALIGNMENTS

RESULT 1
US-09-860-768-2
; Sequence 2, Application US/09860768
; Patent No. 6713289
; GENERAL INFORMATION:
; APPLICANT: Mockel, Bettina
; APPLICANT: Pfefferle, Walter
; APPLICANT: Hermann, Thomas
; APPLICANT: Pohler, Alfred
; APPLICANT: Kalinowski, Jorn
; APPLICANT: Bathe, Brigitte
; TITLE OF INVENTION: New Nucleotide Sequences that Code for the Eno Gene
; FILE REFERENCE: 21123/278404
; CURRENT APPLICATION NUMBER: US/09/860,768
; CURRENT FILING DATE: 2001-05-21
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 2
; LENGTH: 425
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-860-768-2

Query Match	100.0%	Score	2155	DB	4	Length	425
Best Local Similarity	100.0%	Pred. NO.	1.1e-189				
Matches	425	Conservative	0	Mismatches	0	Indels	0
Gaps	0						
Qy	1	VAEIMHVFAREILDSRGNTPTVEAEVFLDDGSHGVAGVPSCAGTGVHAEHLRDCGDRYL	60				
Db	1	VAEIMHVFAREILDSRGNTPTVEAEVFLDDGSHGVAGVPSCAGTGVHAEHLRDCGDRYL	60				
Qy	61	KGVLKAVENNEEIGDELGLADDQRLIDEAMTKLDTANKSRGLGANILGVSMVAKA	120				
Db	61	KGVLKAVENNEEIGDELGLADDQRLIDEAMTKLDTANKSRGLGANILGVSMVAKA	120				
Qy	121	AADSAGLPLFRYIGGPNNAHVLVPMNMIINGGAHSDGVDPQEFMIAPIGAETFEALRN	180				
Db	121	AADSAGLPLFRYIGGPNNAHVLVPMNMIINGGAHSDGVDPQEFMIAPIGAETFEALRN	180				
Qy	181	GAEVYHALKSVIKEKGLSTGLGDEGGFAPSVGSTREALDLIVEAIEKAGTTPGKDIALAL	240				
Db	181	GAEVYHALKSVIKEKGLSTGLGDEGGFAPSVGSTREALDLIVEAIEKAGTTPGKDIALAL	240				
Qy	241	DVASSFFKDGTYHFFGGHSAEMANVYAEVDVPIVSIEDPQDDHGGYTNLTATI	300				
Db	241	DVASSFFKDGTYHFFGGHSAEMANVYAEVDVPIVSIEDPQDDHGGYTNLTATI	300				
Qy	301	GDKVQVIGDDFFVTNPERLKEGIAKKAANSILVKVNOIGTLTTFDADVMHRAGYTSM	360				
Db	301	GDKVQVIGDDFFVTNPERLKEGIAKKAANSILVKVNOIGTLTTFDADVMHRAGYTSM	360				

QY 361 SHRSGETEDTTIADLAVALNCQIKTGAPARSDRVAKYNQLRIEQLLDGAGVYAGRSAP 420
Db |||||
QY 361 SHRSGETEDTTIADLAVALNCQIKTGAPARSDRVAKYNQLRIEQLLDGAGVYAGRSAP 420
Db |||||
QY 421 PRFQ 425
Db |||||
QY 421 PRFQ 425
Db |||||
RESULT 2
US-09-860-768-4
; Sequence 4, Application US/09860768
; Patent No. 6713289
; GENERAL INFORMATION:
; APPLICANT: Mockel, Bettina
; APPLICANT: Pfeifferle, Walter
; APPLICANT: Hermann, Thomas
; APPLICANT: Pohler, Alfred
; APPLICANT: Kalinowski, Jörn
; APPLICANT: Bathe, Brigitte
; TITLE OF INVENTION: New Nucleotide Sequences that Code for the Eno Gene
; FILE REFERENCE: 21123/278404
; CURRENT APPLICATION NUMBER: US/09/860,768
; CURRENT FILING DATE: 2001-05-21
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 4
; LENGTH: 425
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-860-768-4

Query Match 99.8%; Score 2151; DB 4; Length 425;
Best Local Similarity 99.8%; Pred. No. 2.6e-189;
Matches 424; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 VAEIMHVFAREILDNRGNPTVEAEVFLDDGSHGVAGVPSGASTGVHEAHLRGGDRYL 60
Db 1 VAEIMHVFAREILDNRGNPTVEAEVFLDDGSHGVAGVPSGASTGVHEAHLRGGDRYL 60
QY 61 KGVLKAVENNEEIGDELADQRLIDEAMIKLDTANKSRKLGANAILGVSMVAKA 120
Db 61 KGVLKAVENNEEIGDELADQRLIDEAMIKLDTANKSRKLGANAILGVSMVAKA 120
QY 121 AADSAGLPLFRYIGGPNHVLVPVPMNIIINGGAHADSGVDVQEPMTAPIGAETPSEALRN 180
Db 121 AADSAGLPLFRYIGGPNHVLVPVPMNIIINGGAHADSGVDVQEPMTAPIGAETPSEALRN 180
QY 181 GAENVYHALKSVIKKGLSTGLDGGFAPSFGSTREALDLIVEAIEKAGFTPGKDIALAL 240
Db 181 GAENVYHALKSVIKKGLSTGLDGGFAPSFGSTREALDLIVEAIEKAGFTPGKDIALAL 240
QY 241 DVASSEFFKQGYHFEQGQHSAAEMANYAELVDAYPIVSTIEDPLQEDDWEYTNLTATI 300
Db 241 DVASSEFFKQGYHFEQGQHSAAEMANYAELVDAYPIVSTIEDPLQEDDWEYTNLTATI 300
QY 301 GDKVQIVGDDFFVTNPERLKEGIAKKAANSILVKVNOIGTLTETFDVDMAHRAGYTSM 360
Db 301 GDKVQIVGDDFFVTNPERLKEGIAKKAANSILVKVNOIGTLTETFDVDMAHRAGYTSM 360
QY 361 SHRSGETEDTTIADLAVALNCQIKTGAPARSDRVAKYNQLRIEQLLDGAGVYAGRSAP 420
Db 361 SHRSGETEDTTIADLAVALNCQIKTGAPARSDRVAKYNQLRIEQLLDGAGVYAGRSAP 420
QY 421 PRFQ 425
Db |||||
QY 421 PRFQ 425
Db |||||
RESULT 3
US-09-902-540-12060
; Sequence 12060, Application US/09902540

; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 12060
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-12060
Query Match 63.2%; Score 1361.5; DB 4; Length 432;
Best Local Similarity 64.0%; Pred. No. 1.1e-116;
Matches 272; Conservative 58; Mismatches 90; Indels 5; Gaps 2;
QY 1 VAEIMHVFAREILDNRGNPTVEAEVFLDDGSHGVAGVPSGASTGVHEAHLRGG-DRYL 59
Db 1 MTEISQILAREVLDSRGNPTVEAEVQLAGARGRAAVPSGASTGHEAIELRDGDKHRYL 60
QY 60 KGVLKAVENNEEIGDELADQRLIDEAMIKLDTANKSRKLGANAILGVSMVAK 119
Db 61 KGVLKAVENNEEIGDELADQRLIDEAMIKLDTANKSRKLGANAILGVSMVAK 120
QY 120 AADSAGLPLFRYIGGPNHVLVPVPMNIIINGGAHADSGVDVQEPMTAPIGAETPSEALRN 179
Db 120 AADSAGLPLFRYIGGPNHVLVPVPMNIIINGGAHADSGVDVQEPMTAPIGAETPSEALRN 180
QY 180 NGAEVYHALKSVIKKGLSTGLDGGFAPSFGSTREALDLIVEAIEKAGFTPGKDIALA 239
Db 180 NGAEVYHALKSVIKKGLSTGLDGGFAPSFGSTREALDLIVEAIEKAGFTPGKDIALA 240
QY 240 DVASSEFFKQGYHFEQGQHSAAEMANYAELVDAYPIVSTIEDPLQEDDWEYTN 295
Db 240 DVASSEFFKQGYHFEQGQHSAAEMANYAELVDAYPIVSTIEDPLQEDDWEYTN 300
QY 296 LITATIGDKVQIVGDDFFVTNPERLKEGIAKKAANSILVKVNOIGTLTETFDVDMAHRA 355
Db 296 LITATIGDKVQIVGDDFFVTNPERLKEGIAKKAANSILVKVNOIGTLTETFDVDMAHRA 360
QY 301 LITATIGDKVQIVGDDFFVTNPERLKEGIAKKAANSILVKVNOIGTLTETFDVDMAHRA 360
Db 301 LITATIGDKVQIVGDDFFVTNPERLKEGIAKKAANSILVKVNOIGTLTETFDVDMAHRA 415
QY 356 YTSVMSHRSGETEDTTIADLAVALNCQIKTGAPARSDRVAKYNQLRIEQLLDGAGVYA 415
Db 356 YTSVMSHRSGETEDTTIADLAVALNCQIKTGAPARSDRVAKYNQLRIEQLLDGAGVYA 420
QY 416 GRSAP 420
Db |||||
QY 421 GRSAP 425
Db |||||
RESULT 4
US-09-710-279-2222
; Sequence 2222, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: P348008
; CURRENT APPLICATION NUMBER: US/09/710,279
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2222
; LENGTH: 434
; TYPE: PRT

```

; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
US-09-710-279-2222

Query Match      60.4%; Score 1302.5; DB 4; Length 434;
Best Local Similarity 60.6%; Pred. No. 3e-111;
Matches 258; Conservative 65; Mismatches 94; Indels 9; Gaps 3;

Qy 4 IMHVAREILDSRGNPTVEAEVFLDDGSHGVAGVPSGASTGVHEAHELRDGG-DRYLKGK 62
Db 4 ITDYAREVILDSRGNPTVEAEVFLDDGSHGVAGVPSGASTGVHEAHELRDGG-DRYLKGK 63

Qy 63 VLKAVENNEBIGDELAGLEAD--DQRLIDAMIKLDGTANKSRIGANAILGVSMVAKA 120
Db 64 VTKAVENNVNEMIAPEIVEGEFSLVDQVSDIKMWIQLDGTNKGKLGANAILGVSIIVARA 123

Qy 121 AADSAGLPLFRYIGGPNHVLVPVPMNIIINGGAHSDVDVQEFMIAPIGIAETSEALRN 180
Db 124 AADLLGQPLYKLGFGNGKQLPVPMMNIIINGGSHSDAPIAFQEFMILPVGAESFKESLRW 183

Qy 181 GAETHYALKSVIKKGLSTGLDGGFAPSPVSGSTREALDLIVEAEKAGFTPGKDIALAL 240
Db 184 GAETHYALKSVIKKGLSTGLDGGFAPSPVSGSTREALDLIVEAEKAGFTPGKDIALAL 240

Qy 241 DVASSEFFKCTGYHF-----EGQHSAAEMANVYAEVLVDAYPIVSIIDPQEDDWEYGT 294
Db 244 DCASSEFYENGVDYTKFEGEGHAKRSAAEQVDVLEELIGKYPITIEDGMDENDWEGWK 303

Qy 295 NLATIGDKVQIVGDDPFVTPNPERLKEGIAKKAANSILVKVNOIGTTLTETFDVDMAHRA 354
Db 304 QLTDRIQDKVQIVGDDPFVTPNPERLKEGIAKKAANSILVKVNOIGTTLTETFDVDMAHRA 354

Qy 355 GYTSMMSHRSGETEDTTIADLAVALNCGQIKTGAPARSDRVAKYNQLLRIEQLLDGAGVY 414
Db 364 GYTAVVSHRSGETEDTTIADLAVATNAGQIKTGSLSRTDRIAKYNQLLRIEDELITYETAKF 423

Qy 415 AGRSAF 420
Db 424 EGKSF 429

RESULT 5
US-09-134-001C-5641
; Sequence 5641, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR FILING DATE: 1997-11-08
; PRIOR FILING DATE: 1997-11-08
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 5641
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-5641

Query Match      60.4%; Score 1302.5; DB 3; Length 440;
Best Local Similarity 60.6%; Pred. No. 3.1e-111;
Matches 258; Conservative 65; Mismatches 94; Indels 9; Gaps 3;

Qy 4 IMHVAREILDSRGNPTVEAEVFLDDGSHGVAGVPSGASTGVHEAHELRDGG-DRYLKGK 62
Db 10 ITDYAREVILDSRGNPTVEAEVFLDDGSHGVAGVPSGASTGVHEAHELRDGG-DRYLKGK 69

```

```

Qy 63 VLKAVENNEBIGDELAGLEAD--DQRLIDAMIKLDGTANKSRIGANAILGVSMVAKA 120
Db 70 VTKAVENNVNEMIAPEIVEGEFSLVDQVSDIKMWIQLDGTNKGKLGANAILGVSIIVARA 129

Qy 121 AADSAGLPLFRYIGGPNHVLVPVPMNIIINGGAHSDVDVQEFMIAPIGIAETSEALRN 180
Db 130 AADLLGQPLYKLGFGNGKQLPVPMMNIIINGGSHSDAPIAFQEFMILPVGAESFKESLRW 189

Qy 181 GAETHYALKSVIKKGLSTGLDGGFAPSPVSGSTREALDLIVEAEKAGFTPGKDIALAL 240
Db 190 GAETHYALKSVIKKGLSTGLDGGFAPSPVSGSTREALDLIVEAEKAGFTPGKDIALAL 249

Qy 241 DVASSEFFKCTGYHF-----EGQHSAAEMANVYAEVLVDAYPIVSIIDPQEDDWEYGT 294
Db 250 DCASSEFYENGVDYTKFEGEGHAKRSAAEQVDVLEELIGKYPITIEDGMDENDWEGWK 309

Qy 295 NLATIGDKVQIVGDDPFVTPNPERLKEGIAKKAANSILVKVNOIGTTLTETFDVDMAHRA 354
Db 310 QLTDRIQDKVQIVGDDPFVTPNPERLKEGIAKKAANSILVKVNOIGTTLTETFDVDMAHRA 369

Qy 355 GYTSMMSHRSGETEDTTIADLAVALNCGQIKTGAPARSDRVAKYNQLLRIEQLLDGAGVY 414
Db 370 GYTAVVSHRSGETEDTTIADLAVATNAGQIKTGSLSRTDRIAKYNQLLRIEDELITYETAKF 429

Qy 415 AGRSAF 420
Db 430 EGKSF 435

RESULT 6
US-09-252-991A-21020
; Sequence 21020, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 21020
; LENGTH: 444
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-21020

Query Match      59.5%; Score 1282.5; DB 4; Length 444;
Best Local Similarity 61.1%; Pred. No. 2.1e-109;
Matches 261; Conservative 59; Mismatches 100; Indels 7; Gaps 4;

Qy 1 VAEIMHVAREILDSRGNPTVEAEVFLDDGSHGVAGVPSGASTGVHEAHELRDGG-DRYL 59
Db 16 MAKIVDIKREVLDSRGNPTVEAEVFLDDGSHGVAGVPSGASTGVHEAHELRDGG-DRYL 75

Qy 60 KGVLKAVENNEBIGDELAGLEADQRLIDAMIKLDGTANKSRIGANAILGVSMVAKA 119
Db 76 KGVLKAVENNEBIGDELAGLEADQRLIDAMIKLDGTANKSRIGANAILGVSMVAKA 135

Qy 120 AADSAGLPLFRYI-----GGPNHVLVPVPMNIIINGGAHSDVDVQEFMIAPIGIAETFS 175
Db 136 AAAQAKGVPLIYAHADLNGTTPGQYMPVPMNIIINGGAHSDVDVQEFMIAPIGIAETFS 195

Qy 176 EALRNGAEVYHAKSVIKKGLSTGLDGGFAPSPVSGSTREALDLIVEAEKAGFTPGKD 235
Db 196 EALRNGAEVYHAKSVIKKGLSTGLDGGFAPSPVSGSTREALDLIVEAEKAGFTPGKD 255

Qy 236 IALALDVASSEFFKCTGYHFEG-QG-HSAAEMANVYAEVLVDAYPIVSIIDPQEDDWEY 293
Db 236 IALALDVASSEFFKCTGYHFEG-QG-HSAAEMANVYAEVLVDAYPIVSIIDPQEDDWEY 293

```

```
Db      256  VTLALDCASSEFFKDGKYDLEGEKVFDAAGFADYLAGLTQRYPIISIEDGMDESQWAGW 315
Qy      294  TNLTATIGDKVQIVGDDFFVFNTPERLKEGIAKAANSILVKVNOIGTLTTFDAVDMAHR 353
Db      316  KGLTDKIGAKVQVGGDLFVTNTKILKEGIEKGIGNSILKFNQIGSLTETLEAIQWAKA 375
Qy      354  AGYTSMMHSRSGEDEDTTIADLAVALNCGOIKTGAPARSORVAKYNOLLRIEQLLDGAGV 413
Db      376  AGTAVTISHRSGETEDTTIADLAVGTAAQIKTGSCLCRDRVSKYNOLLRIEQLGAKAP 435
Qy      414  YAGRSAP 420
Db      436  YRGRAEF 442

RESULT 7
US-09-830-217-22
; Sequence 22, Application US/09830217
; Patent No. 6521441
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Staphylococcus aureus Genes and Polypeptides
; FILE REFERENCE: PB461PCT
; CURRENT APPLICATION NUMBER: US/09/830,217
; CURRENT FILING DATE: 2001-04-24
; PRIOR APPLICATION NUMBER: PCT/US99/06199
; PRIOR FILING DATE: 1999-03-18
; PRIOR APPLICATION NUMBER: 60/080,296
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/084,674
; PRIOR FILING DATE: 1998-05-07
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 22
; LENGTH: 434
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-830-217-22

Query Match      59.5%; Score 1281.5; DB 4; Length 434;
Best Local Similarity 59.1%; Pred. No. 2.6e-109;
Matches 253; Conservative 68; Mismatches 94; Indels 13; Gaps 4;

Qy      4  IMHVFAREILDSRGNTVEAEVFLDDGSHGVAGVPSCASTGVGHEAHELRODGG-DRYLKKG 62
Db      4  ITDVAREVLDSRGNTVEVEVLTSAGFGRALVPSCASTGEHEAHELRODGDKSRYLKGK 63
Qy      63  VLKAVENVNBEIGDELAGEAD----DQRLIDEAMIKLDGTANKSRILGANAILGVSMVA 118
Db      64  VTRAVENVNBEIIAPEI--IEGEFSLVDQVSDIKMMIALDGTNPKGKLGANAILGVSTAVA 121
Qy      119  KAAADSAGLPLFYIGGPNNAHLVPVPMNIIINGGAHADSGVDVQEFMIAPIGAETTFSEAL 178
Db      122  RAAADLLGQPLYKYLGFNGKQLPVPMMNIIINGGSHSDAPIAFOEFMILPVGATTFKESL 181
Qy      179  RNGAEVYHALKSVIKKGLSTGLDEGGFAPSVGSTREALDLIVEATEKAGFTPGKDIAL 238
Db      182  RWGTEIFHNLSILSQRLGLETAVDEGGFAPKPEGTEDAVETIIQALEAAGYKPGEEVFL 241
Qy      239  ALDVASSEFFKDGTYHF-----EGGQHSAAEMANVYAEVLVDAYPIVISIEDPLQEDDWE 292
Db      242  GFDCASEFYENGVDYKPEGEHGAKRTAAEQVDYLEQLVDKYPITIEDGMDENDWDG 301
Qy      293  YTNLTATIGDKVQIVGDDFFVFNTPERLKEGIAKAANSILVKVNOIGTLTTFDAVDMAH 352
Db      302  WKQLTERIGRVQLVGGDLFVTNTEILAKGIENGIGNSILIKVNOIGTLTTFDAIEMAQ 361
Qy      353  RAGYTSMMHSRSGEDEDTTIADLAVALNCGOIKTGAPARSORVAKYNOLLRIEQLLDGAG 412
Db      362  KAGYTAVVSHRSGETEDTTIADIAVATNAGQIKTGSLSRTDRIAKYNOLLRIEDELPEFA 421
Qy      413  YAGRSAP 420
Db      436  YRGRAEF 442

RESULT 9
US-09-583-110-3817
; Sequence 3817, Application US/09583110
; Patent No. 6699703
; GENERAL INFORMATION:
```

```
Db      422  KYDGIKSF 429

RESULT 8
US-10-278-946-22
; Sequence 22, Application US/10278946
; Patent No. 6821754
; GENERAL INFORMATION:
; APPLICANT: Simpson et al.
; TITLE OF INVENTION: Staphylococcus aureus Genes and Polypeptides
; FILE REFERENCE: PB461USD1
; CURRENT APPLICATION NUMBER: US/10/278,946
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: US 09/830,217
; PRIOR FILING DATE: 2001-04-24
; PRIOR APPLICATION NUMBER: PCT/US99/06199
; PRIOR FILING DATE: 1999-03-18
; PRIOR APPLICATION NUMBER: US 60/078,682
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: US 60/080,296
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: US 60/084,674
; PRIOR FILING DATE: 1998-05-07
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 22
; LENGTH: 434
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-10-278-946-22

Query Match      59.5%; Score 1281.5; DB 4; Length 434;
Best Local Similarity 59.1%; Pred. No. 2.6e-109;
Matches 253; Conservative 68; Mismatches 94; Indels 13; Gaps 4;

Qy      4  IMHVFAREILDSRGNTVEAEVFLDDGSHGVAGVPSCASTGVGHEAHELRODGG-DRYLKKG 62
Db      4  ITDVAREVLDSRGNTVEVEVLTSAGFGRALVPSCASTGEHEAHELRODGDKSRYLKGK 63
Qy      63  VLKAVENVNBEIGDELAGEAD----DQRLIDEAMIKLDGTANKSRILGANAILGVSMVA 118
Db      64  VTRAVENVNBEIIAPEI--IEGEFSLVDQVSDIKMMIALDGTNPKGKLGANAILGVSTAVA 121
Qy      119  KAAADSAGLPLFYIGGPNNAHLVPVPMNIIINGGAHADSGVDVQEFMIAPIGAETTFSEAL 178
Db      122  RAAADLLGQPLYKYLGFNGKQLPVPMMNIIINGGSHSDAPIAFOEFMILPVGATTFKESL 181
Qy      179  RNGAEVYHALKSVIKKGLSTGLDEGGFAPSVGSTREALDLIVEATEKAGFTPGKDIAL 238
Db      182  RWGTEIFHNLSILSQRLGLETAVDEGGFAPKPEGTEDAVETIIQALEAAGYKPGEEVFL 241
Qy      239  ALDVASSEFFKDGTYHF-----EGGQHSAAEMANVYAEVLVDAYPIVISIEDPLQEDDWE 292
Db      242  GFDCASEFYENGVDYKPEGEHGAKRTAAEQVDYLEQLVDKYPITIEDGMDENDWDG 301
Qy      293  YTNLTATIGDKVQIVGDDFFVFNTPERLKEGIAKAANSILVKVNOIGTLTTFDAVDMAH 352
Db      302  WKQLTERIGRVQLVGGDLFVTNTEILAKGIENGIGNSILIKVNOIGTLTTFDAIEMAQ 361
Qy      353  RAGYTSMMHSRSGEDEDTTIADLAVALNCGOIKTGAPARSORVAKYNOLLRIEQLLDGAG 412
Db      362  KAGYTAVVSHRSGETEDTTIADIAVATNAGQIKTGSLSRTDRIAKYNOLLRIEDELPEFA 421
Qy      413  YAGRSAP 420
Db      422  KYDGIKSF 429

RESULT 9
US-09-583-110-3817
; Sequence 3817, Application US/09583110
; Patent No. 6699703
; GENERAL INFORMATION:
```

```

; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; FILE REFERENCE: PATH00-07A
; CURRENT APPLICATION NUMBER: US/09/583,110
; CURRENT FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/107,433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 3817
; LENGTH: 434
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-583-110-3817

```

```

Query Match      58.3%; Score 1256.5; DB 4; Length 434;
Best Local Similarity 58.5%; Pred. No. 5.1e-107;
Matches 250; Conservative 65; Mismatches 101; Indels 11; Gaps 4;

Qy 4 IMHVAREILDSRGNPTVEAEVFLDDSGHGVAGVPSGASTGVHAEHLRDGG-DRYLGKG 62
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4 ITDYAREVLDSRGNPTLEVEVYTESGAFGRGMVPSGASTGHEAEVLRDGDKSRYGGLG 63
Qy 63 VLKAVENNEIEIGDELADGADDOORLIDEAMIKLDGTANKSRGANAILGVSMVAKAAA 122
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
64 TOKAVDNVNNIIABAIIGYDVRDQQAIDRAMIALDGTNPKGKLGANAILGVSIARAAA 123
Qy 123 DSAGLPLFRYIGGNHVLVPMNNIINGGAHDSGVDOEFMIAPIGAEFSEALRNGA 182
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
124 DYLEIPLYSLGGNTKVLPTPMNNIINGSGSHDAPAFQEFMLIPVGATPKREALRYGA 183
Qy 183 EVYHALKSVIKEGLSTGLDGEFGAPSVGSTREALDLIVEAEKAGFTPKDIALALDV 242
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
184 EIFHALKILKSLRGLETAVGDEGGFAPRFEGTGVETILAAIEAAGVVPKGVIFGDC 243
Qy 243 ASSEFFK-----DGTYHFEF--GQHSAAEMANVYAEVDAYDIVSIEDPLQEDDWEY 293
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
244 ASSEFYDKERKVDYDT-KFEFGAGAAVTSAEQIDYLEELVKNKYPITIEDGMDNDWDGW 302
Qy 294 TNLATIGDKVQIVGDDFFVTNPERLKEGIAKKAANSILVKVNOIGTLTETFDVDMWHR 353
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
303 KALTERLGKVKVVGDDFFVTNTDYLRGQIEGANSILIKVNOIGTLTETFEAIEAKE 362
Qy 354 AGYTSMMSHRSGETETTTIADLAVALNCQIKTGAPARSDRVAKYNQLLRIEQLLDAGV 413
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
363 AGYTAVVVSHRSGETETSTIADIAVATNAGQIKTGLSRTDRIAKYNQLLRIEDQLGEVAE 422
Qy 414 YAGRSF 420
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
423 YRLKSF 429

```

```

RESULT 10
US-09-107-532A-3810
; Sequence 3810, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESS: ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; STREET: 100 Beaver Street
; CITY: Walcham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:

```

```

; MEDIUM TYPE: CD/ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 3810:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 398 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...398
; SEQUENCE DESCRIPTION: SEQ ID NO: 3810:
US-09-107-532A-3810

```

```

Query Match      57.7%; Score 1244; DB 4; Length 398;
Best Local Similarity 62.3%; Pred. No. 6.3e-106;
Matches 243; Conservative 58; Mismatches 83; Indels 6; Gaps 3;

Qy 37 VPSGASTGVHAEHLRDGGD-RYLKGVKAVENNEIEIGDELADGADDOORLIDEAMIK 95
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3 VPSGASTGEYAEVLRDGGDKARYGGKGVTKAVDNVNNIIABAIIGYDVRDQMAIDKAMIA 62
Qy 96 LDGTANKSRGANAILGVSMVAKAADSAGLPLFRYIGGNHVLVPMNNIINGGAHA 155
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
63 LDGTNPKGKLGANAILGVSIARAAADYLEVPLHYLGGFNTKVLPTPMNNIINGSGSHA 122
Qy 156 DSGVQVQEFMFIAPIGAEFTSEALRNGAEVYHALKSVIKEGLSTGLDGEFGAPSVGSTR 215
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
123 DNSIDFQEFMIMPVGAPTFKREALRMGAEPHALASILKARGLATSVGDEGFAFNLGSNE 182
Qy 216 EALDLIVEAEKAGFTPKDIALALDVASSEFF--KDGTYHFEF--GQHSAAEMANVYA 270
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
183 EGFEVIEAEKAGYVPGKDVVLAMDAAASSEFVDEKGVVYLADSGEGEKTDEMIFYE 242
Qy 271 ELVDYPIVSIEDPLEQEDDWEYTNLTATIGDKVQIVGDDFFVTNPERLKEGIAKKAANS 330
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
243 ELVSKYPIISIEDGLDENDWDGFKLTDVLGDKVQLVGDDLVFVNTQKLSGIEKGANS 302
Qy 331 ILVKVNOIGTLTETFDVDMWHRAGYTSMMSHRSGETETTTIADLAVALNCQIKTGAPA 390
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
303 ILIKVNOIGTLTETFEAIEAKEAGYTAVVVSHRSGETETSTIADIAVATNAGQIKTGLSLS 362
Qy 391 RSDRVAKYNQLLRIEQLLDAGVYAGRSF 420
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
363 RTDRIAKYNQLLRIEDQLGEVAEYKGLKSF 392

```

```

RESULT 11
US-09-328-352-7269
; Sequence 7269, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.

```

```

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 7269
; LENGTH: 444
; TYPE: PRN
; ORGANISM: Acinetobacter baumannii
US-09-328-352-7269

```

```

Query Match      57.2%; Score 1232.5; DB 4; Length 444;
Best Local Similarity 59.3%; Pred. No. 8.5e-105;
Matches 253; Conservative 65; Mismatches 100; Indels 9; Gaps 5;

QY 1 VAEIMHVPARILSRGNPTVEAEVFLDDGSHGVAGVPSGASTGVHAEHLRDGG-DRYL 59
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
16 MSQIVDIRAREILSRGNPTIEADVILESGVVGACAPSGASTGREALRLDGDKSYL 75
QY 60 KGVLKAVENVEEIGDELGLAGLEADDORLIDEAMIKLDGTANKSRIGANAILGVSMVAK 119
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
76 KGVRTAVQNVNSIHELHVGQSFVEQXALDEKMAFDGTENKSKLGANATLAVSLAAH 135
QY 120 AAASAGLPLRYIG---GPNHVLVPMNIIINGGAHDSGVQVDFMIAPIGAETFS 176
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
136 AAAAEQKLPLFQYIANLRGQTTLTMPVPMNIIINGGAHADNTVDIQSFMIPEVGTSPAE 195
QY 177 ALRGAEVYHAKSVIKEKGLSTGLDGGFAPSVGSTREALDLIVEAIEKAGTTPKDI 236
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
196 ALRGAEVHSLKSVLKQGLTAVGDEGGFAPNLRNEEAITVILQAEIQTYKAGSDI 255
QY 237 ALALDVASSEPFKDGTYHFEQ-GQHS--AARMANVYAEVLVDAYPIVISEDPLQEDDWEY 293
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
256 MALDCCASSEPFYKNGYILEGKNKSTSNQFADYLAGLVKQYPIISIEGLDESDEGW 315
QY 294 TNLTTATGDKVQIVGDDFFVTNPERLKEGIAKKAANSILVKVNOIGTLTTFDAVMAHR 353
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
316 SYLTSILGDKIQLVGDDLFTNPKILQKIDGIDKGVNSILIKYNOIGTLTDLAIYLA 375
QY 354 AGYTSMSHRSGETEDTTIADLAVALNCQIKTGAPARSDRVAKYNOLLRIEQLLDAGV 413
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
376 NGYTTVISHRSGETEDTTIADLAVGTAAGIKTGLSLCRSDRVSKYNOLLRIEEL--TKAV 433
QY 414 YAGRSF 420
Db : : : :
434 YRGKAEF 440

```

```

RESULT 12
US-08-847-065-19
; Sequence 19, Application US/08847065
; Patent No. 6245335
; GENERAL INFORMATION:
; APPLICANT: Masure, H. Robert
; APPLICANT: Rosenow, Carsten I.
; APPLICANT: Tuomanen, Elaine
; APPLICANT: Wizemann, Theresa M.
; TITLE OF INVENTION: CHOLINE BINDING PROTEINS FOR
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David A. Jackson, Esq.
; STREET: 411 Hackensack Ave, Continental Plaza, 4th
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

```

```

; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/847,065
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-158 ..
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 420 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: B. subtilis
US-08-847-065-19

Query Match      56.9%; Score 1226.5; DB 3; Length 420;
Best Local Similarity 58.7%; Pred. No. 2.8e-104;
Matches 244; Conservative 61; Mismatches 100; Indels 11; Gaps 4;

QY 15 SRGNPTVEAEVFLDDGSHGVAGVPSGASTGVHAEHLRDGG-DRYLKGVLKAVENVEE 73
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1 SRGNPTLEVEVYTESGAFGRGMVPSGASTGEHAEVLRDGDKSRYGGLGTQKAVDNVNI 60
QY 74 IGDELAGLEADDORLIDEAMIKLDGTANKSRIGANAILGVSMVAKAAADSAGLPLFYI 133
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
61 IADSIIGYDVRDQQAIDRAMIALDGTNPKGLGANAILGVSI AVARAAADYLEIPLYSYL 120
QY 134 GGNHVLVPMNIIINGGAHDSGVQVDFMIAPIGAETFSALRNGAEVYHAKSVIK 193
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
121 GGNTKVLPTPMNIIINGSHSDAPAFQBFMLPVGAPTFKETTRYGAEIFHALKKILK 180
QY 194 EKLSTGLDGGFAPSVGSTREALDLIVEAIEKAGTFKGDIALDALDVASSEFFK--- 249
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
181 SRGLETAVGDEGGFAPRFEGETDGVETILAAIEAGYVPGKDFLGFDCASSEFYDKERK 240
QY 250 --DGYHFEQ---GQHSAAEMANVYAEVLVDAYPIVISEDPLQEDDWEYTNLTATGDKV 304
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
241 VYDYT-KFEGEGAAVRTSAEQIDYLEELVKNYPIITIEDGMDENDWDGKALTERLGKV 299
QY 305 QIVGDDFFVTNPERLKEGIAKKAANSILVKVNOIGTLTTFDAVMAHRAGYTSMMSHRS 364
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
300 QLVGDDFFVTNTDYLARGIQKGAANSILIKVNOIGTLTTFEAEIEMAKAGYTAVVSHRS 359
QY 365 GETEDTTIADLAVALNCQIKTGAPARSDRVAKYNOLLRIEQLLDAGVYAGRSF 420
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
360 GETEDSTIADIAIATNAGQIKTGSLSRTDRIAKYNOLLRIEQLDGEVAGYGLKSP 415

```

```

RESULT 13
US-09-829-382-19
; Sequence 19, Application US/09829382
; Patent No. 6784164
; GENERAL INFORMATION:
; APPLICANT: Masure, H. Robert
; APPLICANT: Rosenow, Carsten I.
; APPLICANT: Tuomanen, Elaine
; APPLICANT: Wizemann, Theresa M.
; TITLE OF INVENTION: CHOLINE BINDING PROTEINS FOR
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David A. Jackson, Esq.
; STREET: 411 Hackensack Ave, Continental Plaza, 4th

```

; Floor
 ; CITY: Hackensack
 ; STATE: New Jersey
 ; COUNTRY: USA
 ; ZIP: 07601
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/829,382
 ; FILING DATE: 09-Apr-2001
 ; CLASSIFICATION: <Unknown>
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/847,065
 ; FILING DATE: <Unknown>
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Jackson Esq., David A.
 ; REGISTRATION NUMBER: 26,742
 ; REFERENCE/DOCKET NUMBER: 600-1-158 ..
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 201-487-5800
 ; TELEFAX: 201-343-1684
 ; INFORMATION FOR SEQ ID NO: 19:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 420 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: <Unknown>
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; HYPOTHETICAL: NO
 ; ANTI-SENSE: NO
 ; ORIGINAL SOURCE:
 ; ORGANISM: S. pneumoniae
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 19:

Query Match 56.9%; Score 1226.5; DB 4; Length 420;
 Best Local Similarity 58.7%; Pred. No. 2.8e-104;
 Matches 244; Conservative 61; Mismatches 100; Indels 11; Gaps 4;
 Qy 15 SRGNPTVEAEVFLDDGSHGVAGVPSGASTGVHEAHLRDGG-DRYLGKGVILKAVENVNEE 73
 Db 1 SRGNPTLEVEVYTSAGFRCGMVPSGASTGHEAHLRDGKSRYGGLGTQKAVDNVNI 60
 Qy 74 IGDELGLAADDQRLIDEAMIKLDGTANKSRILGANAILGVSMVAKAAGSAGLPLFRYI 133
 Db 61 IADSIIGYDVDRDQAIDRAMTALDGTNPKGLGANAILGVSIARAAADYLEIPLSYL 120
 Qy 134 GGPNAHVLVPMNNIINGGAHDSGVVDQVFEMPIAGAEITSEALRNGAEVYHALKSVIK 193
 Db 121 GGFNTKVLPTPMNNIINGGSHSDAPIAQFQEMILPVGAPTFRYLRYGAEIFHALKKILK 180
 Qy 194 EKLSTGLDGGFPAPSVGSTREALDLIVEAIEKAGTFPGKDIALDLVASESEFFK---- 249
 Db 181 SRGLETAGVDSGGFAPREGTEGDEVETILAEAGVYVPGKDFLGFDCASSEFYDKERK 240
 Qy 250 --DGTYHEG---GQHSAAENAVYAEVLVDAYPIVSDPLQEDDWEGYTNLTATIGDKV 304
 Db 241 VYDYT-KFPEGGAAVRTSAEQIDYLELVNKPITIEDGNDENDWDGKALTERLGKKV 299
 Qy 305 QIVGDDPVTNPERLKEGIAKKAANSILVKNQIGTLTETFDVDMHRACTYMMSHRS 364
 Db 300 QLVGDDFPVTNDYLARGIQGAANSILVKNQIGTLTETFEALMAKEAGYTVAVSHRS 359
 Qy 365 GETEDTTIADLAVALNCQOIKTGAPARSDRVAKYNQLRIEQLLDAGVYAGRSF 420
 Db 360 GETEDSTIADIATNAQOIKTGSISRTDRITAKYNQLRIEQLGEVYAEVGLKSF 415

RESULT 14
 US-09-489-039A-10061

; Sequence 10061, Application US/09489039A
 ; Patent No. 6610836
 ; GENERAL INFORMATION:
 ; APPLICANT: Gary Breton et. al
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
 ; FILE REFERENCE: 2709.2004001
 ; CURRENT APPLICATION NUMBER: US/09/489,039A
 ; CURRENT FILING DATE: 2000-01-27
 ; PRIOR APPLICATION NUMBER: US 60/117,747
 ; PRIOR FILING DATE: 1999-01-29
 ; NUMBER OF SEQ ID NOS: 14342
 ; SEQ ID NO 10061
 ; LENGTH: 459
 ; TYPE: PRT
 ; ORGANISM: Klebsiella pneumoniae
 ; US-09-489-039A-10061

Query Match 56.9%; Score 1226; DB 4; Length 459;
 Best Local Similarity 59.3%; Pred. No. 3.5e-104;
 Matches 252; Conservative 55; Mismatches 110; Indels 8; Gaps 3;
 Qy 1 VAEIMHVPAREILDSRGNPTVEAEVFLDDGSHGVAGVPSGASTGVHEAHLRDGG-DRYL 59
 Db 28 MSKIVKVGIRIIDSRCNPTVEAEVHLEGGFVGMAAAPSGASTGSREALELRDGDKSRFL 87
 Qy 60 GKVLKAVENVNEEIGDELGLAADDQRLIDEAMIKLDGTANKSRILGANAILGVSMVAK 119
 Db 88 KGVTKAAVAVNGPIAQAILGKDAQAGIDKIMIDLDGTENKSNFGANAILAVFLANAK 147
 Qy 120 AAADSAGLPLFRYI---GGFNAHVLVPMNNIINGGAHDSGVVDQVFEMPIAGAEITFS 175
 Db 148 AAAASKGLPLYAHIAELNGTPGKYSMPVPMNNIINGGEHADNNVDIQEFMIQPVGAPTLK 207
 Qy 176 BALNGAEVTHALKSVIKEKGLSTGLDGGFAPSVGSTREALDLIVEAIEKAGFTPKD 235
 Db 208 EAVRMGSEVFHHLAKVLKSKGMNTAVDGGYAPNLGNSAEALAVIAEAVKAAGVELGKD 267
 Qy 236 IALALDVASESEFFKDGTYHFEFGQHS---AEMANVYAEVLVDAYPIVSDPLOEDDWE 292
 Db 268 ITLAMDCASSEFYKDGKYLAGEGNKAFTEEFTHLEELTKQYPIVSDGLDESDEW 327
 Qy 293 YTNLTATIGDKVQIVGDDFFVTNPERLKEGIAKKAANSILVKNQIGTLTETFDVDMH 352
 Db 328 FAYQTKVLGDKIQLVGDLLFVTNTKILKEGIEKGIANSILIKFNQIGSLTETLAAIKWAK 387
 Qy 353 RAGYTSMMSHRSGETEDTTIADLAVALNCQOIKTGAPARSDRVAKYNQLRIEQLLDAG 412
 Db 388 DAGYTAIVISHRSGETEDATTIADLAVGTAAGQIKTGSMSRSDRVAKYNQLRIEALGEOA 447
 Qy 413 VYAGR 417
 Db 448 PFNGR 452

RESULT 15
 US-09-540-236-2354
 ; Sequence 2354, Application US/09540236
 ; Patent No. 6673910
 ; GENERAL INFORMATION:
 ; APPLICANT: Gary L. Breton et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATAR
 ; FILE REFERENCE: 2709.2005-001
 ; CURRENT APPLICATION NUMBER: US/09/540,236
 ; CURRENT FILING DATE: 2000-04-04
 ; NUMBER OF SEQ ID NOS: 3840
 ; SEQ ID NO 2354
 ; LENGTH: 445
 ; TYPE: PRT
 ; ORGANISM: M. catarrhalis
 ; US-09-540-236-2354

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 28, 2005, 16:00:03 ; Search time 793.5 Seconds
(without alignments)
11772.340 Million cell updates/sec

Title: US-10-728-947-3

Perfect score: 1578

Sequence: 1 ggctgggatatggtagtt.....ctcaagcagggaacgtgctt 1578

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_16Dec04:*

- 1: Geneseqn1980s:*
- 2: Geneseqn1990s:*
- 3: Geneseqn2000s:*
- 4: Geneseqn2001as:*
- 5: Geneseqn2001bs:*
- 6: Geneseqn2002as:*
- 7: Geneseqn2002bs:*
- 8: Geneseqn2003as:*
- 9: Geneseqn2003bs:*
- 10: Geneseqn2003cs:*
- 11: Geneseqn2003ds:*
- 12: Geneseqn2004as:*
- 13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1576.4	99.9	1578	4	Aaf61696 C. glutam
2	1576.4	99.9	349980	5	Aah68527 C. glutami
3	1401.8	88.8	1405	10	Add13322 C. glutam
4	1394.8	88.4	1398	4	Aaf71395 Coryneb
5	1276.4	80.9	1278	8	Acc45311 Coryneb
6	1273.4	80.7	1275	5	Aah66050 C. glutami
7	1273.4	80.7	1299	8	Acc45312 Coryneb
8	1271.6	80.6	1278	8	Acc45314 Modified
9	1268.6	80.4	1299	8	Acc45313 Modified
10	907	57.5	1275	8	Acc45313 Modified
11	678.2	43.0	1269	13	AdS56353 Bacterial
12	647	41.0	1287	8	ACA38079 Prokaryot
13	640.8	40.6	30159	4	Aas59564 Propionib
14	640.8	40.6	30159	8	ACF64493 Propionib
15	632.6	40.1	45190	6	Abx09144 Mycobacte
16	632.6	40.1	110000	4	Continuation (12 o
17	632.6	40.1	110000	4	Continuation (12 o
18	632.4	40.1	1317	8	ACA38642 Prokaryot
19	626.2	39.7	1290	8	ACA40436 Prokaryot
20	583.4	37.0	1344	8	ACA39628 Prokaryot

C	21	583	36.9	349980	6	ABQ81846
C	22	583	36.9	349980	6	ABQ81847
	23	580.4	36.8	1269	13	AD58570 Bacterial
	24	577	36.6	1269	13	AD58570 Bacterial
	25	575.6	36.5	1281	13	AD742160 Bacterial
	26	572.4	36.3	1275	13	AD5859832 Bacterial
	27	569	36.1	1260	13	AD5859832 Bacterial
	28	569	36.1	1269	13	AD5859832 Bacterial
	29	568.2	36.0	1278	13	AD743135 Bacterial
	30	567.4	36.0	1272	13	AD743135 Bacterial
	31	561.8	35.6	1272	13	AD585775 Bacterial
	32	558.6	35.4	1278	13	AD746979 Bacterial
	33	550.4	34.9	1290	13	AD744740 Bacterial
	34	542	34.3	1281	8	ACA23750
	35	532	33.7	1335	11	ABD05845 Pseudomon
	36	531.4	33.7	1347	11	ABD05554 Pseudomon
	37	530.2	33.6	1290	4	AAS54197 Pseudomon
	38	530.2	33.6	1290	8	ACA42456 Prokaryot
	39	529.2	33.5	1281	13	AD746255 Bacterial
	40	528	33.5	1287	13	AD56507 Bacterial
	41	522.8	33.1	1278	13	AD549945 Bacterial
	42	522.8	33.1	1281	8	ACA25228 Prokaryot
	43	522.8	33.1	1281	13	AD585460 Bacterial
	44	515.8	32.7	1690	12	ADO59752 B. subtil
	45	513.6	32.5	1404	8	ACA45728 Prokaryot

ALIGNMENTS

RESULT 1

AAF61696
ID AAF61696 standard; DNA; 1578 BP.

XX AAF61696;

XX 12-JUL-2001 (first entry)

XX C. glutamicum enolase encoding DNA.

XX Enolase; fermentation; L-amino acid; L-lysine; coryneform; eno gene;

XX medicine; animal feed supplement; ds.

XX Corynebacterium glutamicum.

XX Key Location/Qualifiers

XX CDS 151..1428

XX /*tag= a

XX /product= "Enolase"

XX EP1090998-A1.

XX 11-APR-2001.

XX 29-SEP-2000; 2000EP-00121158.

XX 05-OCT-1999; 99DE-01047791.

XX (DEGS) DEGUSSA-HUELS AG.

XX Moeckel B, Pfeifferle W, Hermann T, Puehler A, Kalinowski J;

XX Bathe B;

XX WPI; 2001-292928/31.

XX P-PSDB; AAB70882.

XX New enolase gene from coryneform bacteria, used to prepare transformants

XX with increased synthesis of amino acids, particularly lysine.

XX Claim 4; Page 12-14; 25pp; German.

XX This invention describes a novel isolated nucleic acid (I) from

XX coryneform bacteria which is used in a method for fermentative production

CC of L-amino acids, especially L-lysine, by fermenting a lysine-producing
 CC coryneform in which the eno gene has been amplified, and isolating amino
 CC acids that have accumulated in the medium or cells. (1), which encodes an
 CC enolase, is used to transform coryneforms for production of L-amino
 CC acids, specifically lysine which is used in medicine and particularly as
 CC animal feed supplement. It may also be used as probes and primers for
 CC isolating related sequences. Overexpression of (1) improves production of
 CC amino acids, especially of L-lysine. This sequence encodes the enolase
 CC protein described in the method of the invention
 XX
 SQ Sequence 1578 BP; 351 A; 461 C; 424 G; 342 T; 0 U; 0 Other;

Query Match 99.98; Score 1576.4; DB 4; Length 1578;
 Best Local Similarity 99.98; Pred. No. 0;
 Matches 1577; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGCTGGGGATATGGGTAGTTTTCGCCACTAATTTCAACTGATTCGCTCATCGAAACAAGA 60
 DB 1 GGCTGGGGATATGGGTAGTTTTCGCCACTAATTTCAACTGATTCGCTCATCGAAACAAGA 60

QY 61 TTCTGTCAACAATTTGGGTAGAGTGAATTAAGACATTTTGAATGATCATCGTGAATTAATCTAG 120
 DB 61 TTCTGTCAACAATTTGGGTAGAGTGAATTAAGACATTTTGAATGATCATCGTGAATTAATCTAG 120

QY 121 TTAGTCTCCCAAGTTGGCATAGGAGGCCACAGTGGCTGAAATCATGACGTTATTCGCTCGC 180
 DB 121 TTAGTCTCCCAAGTTGGCATAGGAGGCCACAGTGGCTGAAATCATGACGTTATTCGCTCGC 180

QY 181 GAAATTCCTGACCTCCCGGGTAAACCAACCGTTCAGGACAGAGGTTTTCCTGGATGACGGT 240
 DB 181 GAAATTCCTGACCTCCCGGGTAAACCAACCGTTCAGGACAGAGGTTTTCCTGGATGACGGT 240

QY 241 TCCACAGGTTCGAGGTGTTTCATCCGGCGCATCCACCGGCGTCCACGAGGTCATGAG 300
 DB 241 TCCACAGGTTCGAGGTGTTTCATCCGGCGCATCCACCGGCGTCCACGAGGTCATGAG 300

QY 301 CTGGTTCAGCGTGGCGATCGTACTCGGCAAGGGCGTTTGAAGGAGTTGAAGAGTCC 360
 DB 301 CTGGTTCAGCGTGGCGATCGTACTCGGCAAGGGCGTTTGAAGGAGTTGAAGAGTCC 360

QY 361 AACGAAGAAATCGGCGACGAGTCCGCTGGCTAGAGGCTGACGATCAGCGCTCATCGAC 420
 DB 361 AACGAAGAAATCGGCGACGAGTCCGCTGGCTAGAGGCTGACGATCAGCGCTCATCGAC 420

QY 421 GAAGCAATGATCAAGCTTGATGGCAACCGCAACAAAGTCCCGCTGGGTGCAACGCAATC 480
 DB 421 GAAGCAATGATCAAGCTTGATGGCAACCGCAACAAAGTCCCGCTGGGTGCAACGCAATC 480

QY 481 CTTGGTGTTCATGGCTGTTGCAAGGCTGCTGCTGATTCGCGAGGCTCCCACTGTC 540
 DB 481 CTTGGTGTTCATGGCTGTTGCAAGGCTGCTGCTGATTCGCGAGGCTCCCACTGTC 540

QY 541 CGCTACATCGGTGACCAACGACAGTTCCTCCAGTTTCCAATGATGAACATCATCAAC 600
 DB 541 CGCTACATCGGTGACCAACGACAGTTCCTCCAGTTTCCAATGATGAACATCATCAAC 600

QY 601 GGTGGCGTCAAGCTGATCGGTTGATGAGTTCAGGAAATTCATGATCGCTCCCAATCGGT 660
 DB 601 GGTGGCGTCAAGCTGATCGGTTGATGAGTTCAGGAAATTCATGATCGCTCCCAATCGGT 660

QY 661 GCAGAGACCTTCTCTGAGGCTCTCGCAACGGCGGAGGCTTACACGACGACTGAAGTCC 720
 DB 661 GCAGAGACCTTCTCTGAGGCTCTCGCAACGGCGGAGGCTTACACGACGACTGAAGTCC 720

QY 721 GTCATCAAGGAAAGGGCTCTCCACCGGACTTGGCGATGAGGGCGCTTCGCTCCCTCC 780
 DB 721 GTCATCAAGGAAAGGGCTCTCCACCGGACTTGGCGATGAGGGCGCTTCGCTCCCTCC 780

QY 781 GTGGGCTCCACCGGTGAGGCTCTTGACCTTAAGGCAATTCGAGAGGCTGGGCTTC 840
 DB 781 GTGGGCTCCACCGGTGAGGCTCTTGACCTTAAGGCAATTCGAGAGGCTGGGCTTC 840

QY 841 ACCCCAGGCAAGGACATCGCTCTTGGCTGAGAGTTCCTCTCTGAGTTCCTCAAGGAC 900

DB 841 ACCCCAGGCAAGGACATCGCTCTTGGCTGAGAGTTCCTCTGAGTTCCTCAAGGAC 900
 QY 901 GGCACCTACCACTTCGAAGGTGCCAGCACTCCGAGCTGAGATGCAAAAGTTTACGCT 960
 DB 901 GGCACCTACCACTTCGAAGGTGCCAGCACTCCGAGCTGAGATGCAAAAGTTTACGCT 960

QY 961 GAGCTGTTGACGCGTACCAATCGTTCATCGAGGACCCACTGCGAGGAGTACTGG 1020
 DB 961 GAGCTGTTGACGCGTACCAATCGTTCATCGAGGACCCACTGCGAGGAGTACTGG 1020

QY 1021 GAGGGTTTACACCAACCTCACCGCAACCATCGGCGCAAGGTTTCAGATCGTTGCGGACGAC 1080
 DB 1021 GAGGGTTTACACCAACCTCACCGCAACCATCGGCGCAAGGTTTCAGATCGTTGCGGACGAC 1080

QY 1081 TTCTTCGTCAACCAACCTGAGCGCTGAAGAGGGGATCGCTAAGAGGCTGCAACTCC 1140
 DB 1081 TTCTTCGTCAACCAACCTGAGCGCTGAAGAGGGGATCGCTAAGAGGCTGCAACTCC 1140

QY 1141 ATCTCTGTTAAGGTGAACCAAGATCGGTACCTCACCGAGACCTTCGACGCTGTCGACATG 1200
 DB 1141 ATCTCTGTTAAGGTGAACCAAGATCGGTACCTCACCGAGACCTTCGACGCTGTCGACATG 1200

QY 1201 GCTCACCGCGCAGGCTACACCTCCATGATGTCCACCGTTCGCGTGAGACCCGAGGACACC 1260
 DB 1201 GCTCACCGCGCAGGCTACACCTCCATGATGTCCACCGTTCGCGTGAGACCCGAGGACACC 1260

QY 1261 ACCATTGCTGACCTCCGAGTTGCACTCAACTGTGGCGAGATCAAGACTGGTGTCTCCAGCA 1320
 DB 1261 ACCATTGCTGACCTCCGAGTTGCACTCAACTGTGGCGAGATCAAGACTGGTGTCTCCAGCA 1320

QY 1321 CGTTCCGACCGTGTCCAAAGTACACCAAGTTCCTCCGATCGAGAGCTGCTTGGCGAC 1380
 DB 1321 CGTTCCGACCGTGTCCAAAGTACACCAAGTTCCTCCGATCGAGAGCTGCTTGGCGAC 1380

QY 1381 GCGGCGGTCTACGACGCTGCGACGATTCCTCCACGCTTTCAGGGCTTAAATAAAGGCGCTT 1440
 DB 1381 GCGGCGGTCTACGACGCTGCGACGATTCCTCCACGCTTTCAGGGCTTAAATAAAGGCGCTT 1440

QY 1441 TTCCGACCGCGGTAACTCAAGGTTCCCGGGCGTCCGCTTACTGTTACTGGTGT 1500
 DB 1441 TTCCGACCGCGGTAACTCAAGGTTCCCGGGCGTCCGCTTACTGTTACTGGTGT 1500

QY 1501 GACTATGATCGAGGATTTATGGCAAGCAAGCAAGAACTCATAAAGGCTTGTTCCTGTCT 1560
 DB 1501 GACTATGATCGAGGATTTATGGCAAGCAAGCAAGAACTCATAAAGGCTTGTTCCTGTCT 1560

QY 1561 CAAGCAGGGAACGTCGCTT 1578
 DB 1561 CAAGCAGGGAACGTCGCTT 1578

RESULT 2
 AAH68527
 ID AAH68527 standard; DNA; 349980 BP.
 XX
 AC AAH68527;
 XX
 DT 26-SEP-2001 (first entry)
 XX
 DE C glutamicum coding sequence fragment SEQ ID NO: 7062.
 XX
 KW Corynebacterium; amino acid synthesis; vitamin; saccharide;
 XX organic acid synthesis; ds.
 OS Corynebacterium glutamicum.
 XX
 PN EF1108790-A2.
 XX
 PD 20-JUN-2001.
 XX
 PF 18-DEC-2000; 2000EP-00127688.
 XX

PR 16-DEC-1999; 99JP-00377484.
 PR 07-APR-2000; 2000JP-00159162.
 PR 03-AUG-2000; 2000JP-00280988.
 XX (KYOW) KYOWA HAKKO KOGYO KK.
 XX Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
 XX Tateishi N, Senoh A, Ikeda M, Ozaki A;
 XX WPI; 2001-376931/40.
 XX Novel polynucleotides derived from Coryneform bacteria, for identifying
 PT mutation point of a gene, measuring expression of a gene, analyzing
 PT expression profile or pattern of a gene and identifying homologous gene.
 XX Disclosure; SEQ ID NO 7062; 246pp + Sequence Listing; English.
 CC The present invention provides a number of nucleotide and protein
 CC sequences from the Coryneform bacterium *Corynebacterium glutamicum*. These
 CC are useful for identifying the mutation point of a gene derived from a
 CC mutant of coryneform bacterium, measuring expression amount and analysing
 CC the expression profile or expression pattern of a gene derived from
 CC Coryneform bacterium, and identifying a homologue of a gene derived from
 CC coryneform bacterium. Coryneform bacteria are useful for producing amino
 CC acids, nucleic acids, vitamins, saccharides and organic acids,
 CC particularly L-lysine. The present sequence is a nucleic acid described
 CC in the exemplification of the invention. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from the European Patent Office
 XX
 SQ Sequence 349980 BP; 79725 A; 90426 C; 98918 G; 80911 T; 0 U; 0 Other;
 Query Match 99.9%; Score 1576.4; DB 5; Length 349980;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1577; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 GGCTGGGGATATGGGTAGTTTCGCCACTATTTCACTGATTCGCTCATCGAAACAAGA 60
 Db 134799 GGCTGGGGATATGGGTAGTTTCGCCACTATTTCACTGATTCGCTCATCGAAACAAGA 134858
 Qy 61 TTCTGTCAACAATTTGGGTGTAGAGCGTATTGAAGACATTTGATCAGTGAATAATTTCTAG 120
 Db 134859 TTCTGTCAACAATTTGGGTGTAGAGCGTATTGAAGACATTTGATCAGTGAATAATTTCTAG 134918
 Qy 121 TTACTCCCAAGTTGGCATAGGAGCCACAGTGGCTGAATCATGCAAGTATTCGCTCGC 180
 Db 134919 TTACTCCCAAGTTGGCATAGGAGCCACAGTGGCTGAATCATGCAAGTATTCGCTCGC 134978
 Qy 181 GAAATTTCTCGACTCCCGCGTAAACCAACCGTCCGAGCAGAGGTTTTCCTGGATGACGGT 240
 Db 134979 GAAATTTCTCGACTCCCGCGTAAACCAACCGTCCGAGCAGAGGTTTTCCTGGATGACGGT 135038
 Qy 241 TCCACCGGTTCGAGGTGTTCCATCCGCGGCATCCACCGCGTCCAGGGCTCATGAG 300
 Db 135039 TCCACCGGTTCGAGGTGTTCCATCCGCGGCATCCACCGCGTCCAGGGCTCATGAG 135098
 Qy 301 CTGCGTACGGTGGCGATCGTACTCTGGGCAAGGGCGTTTGAAGGACAGTTGAAAACGTC 360
 Db 135099 CTGCGTACGGTGGCGATCGTACTCTGGGCAAGGGCGTTTGAAGGACAGTTTGAAGGACGTC 135158
 Qy 361 AACGAAGAAATCGGCGAGCGTCTGGCTAGAGGCTGACGATCAGCGCTCATCGAC 420
 Db 135159 AACGAAGAAATCGGCGAGCGTCTGGCTAGAGGCTGACGATCAGCGCTCATCGAC 135218
 Qy 421 GAAGCAATGATCAAGCTTTGATGGCACCGCCAAAGTCCCGCTGGGTGCAAAACGCAATC 480
 Db 135219 GAAGCAATGATCAAGCTTTGATGGCACCGCCAAAGTCCCGCTGGGTGCAAAACGCAATC 135278
 Qy 481 CTTGGTGTTCATGGCTGTTGCAAAAGCTGCTGCTGATTCGGAGGCTCCCACTGTTC 540
 Db 135279 CTTGGTGTTCATGGCTGTTGCAAAAGCTGCTGCTGATTCGGAGGCTCCCACTGTTC 135338
 Qy 541 CGCTACATCGGTGACCAAGCGCACAGTTCCTCCAGTTCCTCAATGATCAATCATCAAC 600

Db	135339	CGCTACATCGGTGACCAACGACACAGTTCCTCCAGTTCCTCAATGATGAACATCATCAAC	135398
Qy	601	GTTGGCGCTCAGCTGACTCGGTGTTGAGTTCAGGAATTCATGATCGCTCCAAATCGGT	660
Db	135399	GTTGGCGCTCAGCTGACTCGGTGTTGAGTTCAGGAATTCATGATCGCTCCAAATCGGT	135458
Qy	661	GCAGAGACCTTTCTCTGAGGCTCTCCGCAAGCGCGGAGGTCTACACGCACTGAAGTCC	720
Db	135459	GCAGAGACCTTTCTCTGAGGCTCTCCGCAAGCGCGGAGGTCTACACGCACTGAAGTCC	135518
Qy	721	GTCAATCAAGGAAAGGCGCTGTCCACCGGACTTTGGCGATGAGGGCGCTTCGCTCCTTCC	780
Db	135519	GTCAATCAAGGAAAGGCGCTGTCCACCGGACTTTGGCGATGAGGGCGCTTCGCTCCTTCC	135578
Qy	781	GTGGGCTCCACCGCTGAGGCTCTTGACCTTATCGTTAAGCAATCGAGAGGCTGGCTTC	840
Db	135579	GTGGGCTCCACCGCTGAGGCTCTTGACCTTATCGTTAAGCAATCGAGAGGCTGGCTTC	135638
Qy	841	ACCCAGGCAAGGACATCGCTCTTGCTCTGGACCTTGCTCTCTGAGTTCCTCAAGGAC	900
Db	135639	ACCCAGGCAAGGACATCGCTCTTGCTCTGGACCTTGCTCTCTGAGTTCCTCAAGGAC	135698
Qy	901	GGCACTTACCACTTCGAAGGTGGCCAGACTCCCGAGCTGAGATGGCAAAAGTTCAGCT	960
Db	135699	GGCACTTACCACTTCGAAGGTGGCCAGACTCCCGAGCTGAGATGGCAAAAGTTCAGCT	135758
Qy	961	GAGCTCGTTGACCGGTACCAATCGTCTCCATCGAGGACCCACTCGAGGAGATGACTGG	1020
Db	135759	GAGCTCGTTGACCGGTACCAATCGTCTCCATCGAGGACCCACTCGAGGAGATGACTGG	135818
Qy	1021	GAGGTTTACACCACTTCACCGCAACCATCGGCGCAAGGTTTCAGATCGTTGGCGACGAC	1080
Db	135819	GAGGTTTACACCACTTCACCGCAACCATCGGCGCAAGGTTTCAGATCGTTGGCGACGAC	135878
Qy	1081	TTCTTCGTCAACCACTTCGAGCGCTGGAAGGGCATCGCTAAGAGAGGCTGCCAACTCC	1140
Db	135879	TTCTTCGTCAACCACTTCGAGCGCTGGAAGGGCATCGCTAAGAGAGGCTGCCAACTCC	135938
Qy	1141	ATCTCGTTAAGGTGAACAGATCGGTACCTCACCAGACCTTCGAGCGCTGTCGACATG	1200
Db	135939	ATCTCGTTAAGGTGAACAGATCGGTACCTCACCAGACCTTCGAGCGCTGTCGACATG	135998
Qy	1201	GCTCACCGCGAGCTACCTCCATCATGTGCCACCGTTCCCGTGAGACCGAGGACACC	1260
Db	135999	GCTCACCGCGAGCTACCTCCATCATGTGCCACCGTTCCCGTGAGACCGAGGACACC	136058
Qy	1261	ACCATTCGTGACCTTCGAGTTGCACTCAACTGTGGCCAGATCAAGACTGGTGTCCAGCA	1320
Db	136059	ACCATTCGTGACCTTCGAGTTGCACTCAACTGTGGCCAGATCAAGACTGGTGTCCAGCA	136118
Qy	1321	CGTTCCGACCGTGTGCAAAAGTACAACAGCTTCCTCGCATCGAGCAGCTGCTGGCGAC	1380
Db	136119	CGTTCCGACCGTGTGCAAAAGTACAACAGCTTCCTCGCATCGAGCAGCTGCTGGCGAC	136178
Qy	1381	GCAGGCTCTACGAGCTCCGAGCGCATTCACCGCTTCAGGGCTAAATAAAGCGCTT	1440
Db	136179	GCAGGCTCTACGAGCTCCGAGCGCATTCACCGCTTCAGGGCTAAATAAAGCGCTT	136238
Qy	1441	TTGACGCCCGGTAACTCAAGGTTGCGGGCGCTGTTGCTTACTACTGTTACTGTTGT	1500
Db	136239	TTGACGCCCGGTAACTCAAGGTTGCGGGCGCTGTTGCTTACTACTGTTACTGTTGT	136298
Qy	1501	GACTATGATCGAGTATGCGAAAGCAGAGAAAACTCATAAAGGCTTGTCTCTGTCT	1560
Db	136299	GACTATGATCGAGTATGCGAAAGCAGAGAAAACTCATAAAGGCTTGTCTCTGTCT	136358
Qy	1561	CAAGCAGGGAACGTGCTT 1578	
Db	136359	CAAGCAGGGAACGTGCTT 136376	

RESULT 3

QY 1191 TGTGACATGGCTACCCCGGAGGCTACACCTCCATGATGTCCACCGTTCGGTGAGAC 1250
Db 1141 TGTGACATGGCTACCCCGGAGGCTACACCTCCATGATGTCCACCGTTCGGTGAGAC 1200
QY 1251 CGAGGACACCACTTGTGACCTCGCAGTTGCACTCAACTGTGGCCAGATCAAGACTGG 1310
Db 1201 CGAGGACACCACTTGTGACCTCGCAGTTGCACTCAACTGTGGCCAGATCAAGACTGG 1260
QY 1311 TGCTCCAGCAGTTCCGACCGTGTGCGCAAGTACAACCAAGCTTCTCCGCATCGAGCAGCT 1370
Db 1261 TGCTCCAGCAGTTCCGACCGTGTGCGCAAGTACAACCAAGCTTCTCCGCATCGAGCAGCT 1320
QY 1371 GCTTGGCAGCGCGGCTTACGAGGTCGAGCGCATTCACCGTTCAGGCTAAAT 1430
Db 1321 GCTTGGCAGCGCGGCTTACGAGGTCGAGCGCATTCACCGTTCAGGCTAAAT 1380
QY 1431 AAAAGCGCTTTTCGACCGCCGGTAA 1455
Db 1381 AAAAGCGCTTTTCGACCGCCGGTAA 1405

RESULT 4
ID AAF71395 standard; DNA; 1398 BP.
AC AAF71395;
DT 30-APR-2001 (first entry)
DE Corynebacterium glutamicum SMP protein nucleotide sequence SEQ ID NO:71.
KW Corynebacterium glutamicum; carbon metabolism and energy production;
KW SMP protein; sugar metabolism and oxidative phosphorylation protein;
KW fine chemical production; organic acid; proteinogenic amino acid;
KW nonproteinogenic amino acid; purine base; pyrimidine base; nucleoside;
KW nucleotide; lipid; saturated fatty acid; unsaturated fatty acid; diol;
KW carbohydrate; aromatic compound; vitamin; cofactor; polyketide; enzyme;
KW diagnosis; Corynebacterium diphtheriae; evolutionary study; ds.
OS Corynebacterium glutamicum.

PN MO200100844-A2.
XX PD 04-JAN-2001.
XX 23-JUN-2000; 2000WO-IB000943.
XX 25-JUN-1999; 99US-0141031P.
PR 08-JUL-1999; 99DE-01031412.
PR 08-JUL-1999; 99DE-01031413.
PR 08-JUL-1999; 99DE-01031419.
PR 08-JUL-1999; 99DE-01031420.
PR 08-JUL-1999; 99DE-01031424.
PR 08-JUL-1999; 99DE-01031428.
PR 08-JUL-1999; 99DE-01031431.
PR 08-JUL-1999; 99DE-01031433.
PR 08-JUL-1999; 99DE-01031434.
PR 08-JUL-1999; 99DE-01031510.
PR 08-JUL-1999; 99DE-01031562.
PR 08-JUL-1999; 99DE-01031634.
PR 09-JUL-1999; 99DE-01032180.
PR 09-JUL-1999; 99DE-01032227.
PR 09-JUL-1999; 99DE-01032230.
PR 09-JUL-1999; 99US-0143208P.
PR 14-JUL-1999; 99DE-01032924.
PR 14-JUL-1999; 99DE-01032973.
PR 14-JUL-1999; 99DE-01033005.
PR 27-AUG-1999; 99DE-01040765.
PR 31-AUG-1999; 99US-0151572P.
PR 03-SEP-1999; 99DE-01042076.
PR 03-SEP-1999; 99DE-01042079.
PR 03-SEP-1999; 99DE-01042086.
PR 03-SEP-1999; 99DE-01042087.

PR 03-SEP-1999; 99DE-01042088.
PR 03-SEP-1999; 99DE-01042095.
PR 03-SEP-1999; 99DE-01042123.
PR 03-SEP-1999; 99DE-01042125.
XX (BADI) BASF AG.
PA Pompejus M, Kroeger B, Schroeder H, Zelder O, Haberhauer G;
XX WPI: 2001-061975/07.
PI P-PSDB; AAB79278.
DR
XX
XX New isolated Corynebacterium glutamicum nucleic acid encoding a sugar
metabolism and oxidative phosphorylation protein for production or
modulation of production of fine chemicals e.g. amino acids,
carbohydrates or enzymes.
XX Claim 3; Page 242-244; 1246pp; English.
XX AAF71360 to AAF71750 encode the Corynebacterium glutamicum sugar
metabolism and oxidative phosphorylation (SMP) proteins given in AAB79243
to AAB 79633 which are involved in carbon metabolism and energy
production. The C. glutamicum SMP gene can be used in vectors (II) for
expression in host cells and production or modulation of production of
fine chemicals, such as, an organic acid, a proteinogenic or
nonproteinogenic amino acid (preferred), a purine or pyrimidine base, a
nucleoside, a nucleotide, a lipid, a saturated or unsaturated fatty acid,
a diol, a carbohydrate, an aromatic compound, a vitamin, a cofactor, a
polyketide, or an enzyme. The presence of (I) or SMP proteins (III)
encoded by them are used for diagnosing the presence or activity of
Corynebacterium diphtheriae in a subject. (I), (II), (III) or host cells
containing them are used to map genomes of organisms related to C.
glutamicum, identify and localise C. glutamicum sequences of interest, in
evolutionary studies, in determining SMP protein regions required for
function, in modulating SMP protein activity, in modulating the
metabolism of sugars, and in modulating high-energy molecule production
in a cell (i.e. ATP, NADPH)
XX
SQ Sequence 1398 BP; 308 A; 423 C; 376 G; 291 T; 0 U; 0 Other;

Query Match 88.4%; Score 1394.8; DB 4; Length 1398;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1396; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 51 CGAACAAGATTCTGTCACAAATTTGGGTGTAGAGCTGATTGAAGACATTTTATCACGTGA 110
Db 1 CGAACAAGATTCTGTCACAAATTTGGGTGTAGAGCTGATTGAAGACATTTTATCACGTGA 60
QY 111 ATAATTTCTAGTTAGTCTCCCAAGTTGGCATAGGAGGCCACAGTGGCTGAAATTCATGCACGT 170
Db 61 ATAATTTCTAGTTAGTCTCCCAAGTTGGCATAGGAGGCCACAGTGGCTGAAATTCATGCACGT 120
QY 171 ATTGCTCGCGAAATTTCTCGACTCCCGCGGTAAACCCACCGTCGAGGACAGAGTTTTCCT 230
Db 121 ATTGCTCGCGAAATTTCTCGACTCCCGCGGTAAACCCACCGTCGAGGACAGAGTTTTCCT 180
QY 231 GGATGACGGTTCACAGGTGTCCAGTGTTCATCCGCGCATCCACCGGCTCCACGA 290
Db 181 GGATGACGGTTCACAGGTGTTCATCCGCGCATCCACCGGCTCCACGA 240
QY 291 GGCTCATGAGCTCGGTGACGGTGGCGCATCGTACCTGGGCAAGGGCGTTTGAAGGCAGT 350
Db 241 GGCTCATGAGCTCGGTGACGGTGGCGCATCGTACCTGGGCAAGGGCGTTTGAAGGCAGT 300
QY 351 TGAACACGTCAACGAAGAAATTCGCGACAGAGTTCGCTGGCTTAGAGGCTGACGATCAGCG 410
Db 301 TGAACACGTCAACGAAGAAATTCGCGACAGAGTTCGCTGGCTTAGAGGCTGACGATCAGCG 360
QY 411 CCTCATCGAGCAAGCAATGATCAAGCTTGATGCGACCGGCCAACAAAGTCCCGCTGGGTGC 470
Db 361 CCTCATCGAGCAAGCAATGATCAAGCTTGATGCGACCGGCCAACAAAGTCCCGCTGGGTGC 420
QY 471 AAACGCAATCCTTGGTGTTCATGGCTGTGTGCAAGGCTGCTGCTGATTCGCGAGGCCT 530

Db 421 AAAGCAATCTTGGTGTTCATGGCTGTGCAAGGCTGCTGTGATTCGACGGCCT 480
 Qy 531 CCACATGTTCCGCTACATCGGTGGACAAACGACACAGTTCCTTCCAGTTCCTCAATGATGAA 590
 Db 481 CCCACTGTTCCGCTACATCGGTGGACAAACGACACAGTTCCTTCCAGTTCCTCAATGATGAA 540
 Qy 591 CATCATCAACGGTGGCGCTCACGCTGACTCCGCTGCTGAGTTCAGGAATTCATGATCGC 650
 Db 541 CATCATCAACGGTGGCGCTCACGCTGACTCCGCTGCTGAGTTCAGGAATTCATGATCGC 600
 Qy 651 TCCAATCGGTGCAGAGACCTTCTCTGAGGCTCTCCGCAACGGCGCGAGGTCTACCAACGC 710
 Db 601 TCCAATCGGTGCAGAGACCTTCTCTGAGGCTCTCCGCAACGGCGCGAGGTCTACCAACGC 660
 Qy 711 ACTGAAGTCCGTCATCAAGAAAGGGCCTGTCCACCGGACTTGGCGATGAGGCGGCTT 770
 Db 661 ACTGAAGTCCGTCATCAAGAAAGGGCCTGTCCACCGGACTTGGCGATGAGGCGGCTT 720
 Qy 771 CGCTCCTTCCGTCGGCTCCACCGGTGAGGCTCTTGACCTTATCGTTAAGGCAATCGAGAA 830
 Db 721 CGCTCCTTCCGTCGGCTCCACCGGTGAGGCTCTTGACCTTATCGTTAAGGCAATCGAGAA 780
 Qy 831 GGCTGGCTTCAACCCAGGCAAGACATCGCTCTTGCTCTGAGCGTTCCTTCTCTGAGTT 890
 Db 781 GGCTGGCTTCAACCCAGGCAAGACATCGCTCTTGCTCTGAGCGTTCCTTCTCTGAGTT 840
 Qy 891 CTTCAAGGACGGCACCTTACCACTTCGAAGTGTGGCCAGCACTCCGACGCTGAGATGGCAAA 950
 Db 841 CTTCAAGGACGGCACCTTACCACTTCGAAGTGTGGCCAGCACTCCGACGCTGAGATGGCAAA 900
 Qy 951 CGTTAGCTGAGTCTGCTGACCGGTACCAACCTGCTCCATCGTCCATCGAGGACCACTGCGAG 1010
 Db 901 CGTTAGCTGAGTCTGCTGACCGGTACCAACCTGCTCCATCGTCCATCGAGGACCACTGCGAG 960
 Qy 1011 AGATGACTGGAGGGTTTACACCAACCTCAGCGCAACCATCGGCGCAACAGGTTTCAGATCGT 1070
 Db 961 AGATGACTGGAGGGTTTACACCAACCTCAGCGCAACCATCGGCGCAACAGGTTTCAGATCGT 1020
 Qy 1071 TGGCGACGATCTTCTGTCACCAACCTGAGCGCTGAAAGGAGGCGCATCGCTAAGAGGC 1130
 Db 1021 TGGCGACGATCTTCTGTCACCAACCTGAGCGCTGAAAGGAGGCGCATCGCTAAGAGGC 1080
 Qy 1131 TGGCACTCCATCTGCTGAGTGAACGAGTACCGGTACCGCTCACCGAGACCTTCGACGC 1190
 Db 1081 TGGCACTCCATCTGCTGAGTGAACGAGTACCGGTACCGCTCACCGAGACCTTCGACGC 1140
 Qy 1191 TGTGACATGGCTCACCGCGCAGGCTACACCTCCATGATGTCCACCGTTCGCGTGAGAC 1250
 Db 1141 TGTGACATGGCTCACCGCGCAGGCTACACCTCCATGATGTCCACCGTTCGCGTGAGAC 1200
 Qy 1251 CGAGGACACCAACCTGCTGACCTCGAGTGTGCACTCAATGTGGCCAGATCAAGACTGG 1310
 Db 1201 CGAGGACACCAACCTGCTGACCTCGAGTGTGCACTCAATGTGGCCAGATCAAGACTGG 1260
 Qy 1311 TGCTCCAGCAGCTTCGACCGGTGCGAAAGTCAACACGAGTCTCCGATCCGACGAGCT 1370
 Db 1261 TGCTCCAGCAGCTTCGACCGGTGCGAAAGTCAACACGAGTCTCCGATCCGACGAGCT 1320
 Qy 1371 GCTTGGCGAGCCCGGCTGTACGAGGCTCGCAGCGGCTATCCACGCTTTCAGGCGCTAAAT 1430
 Db 1321 GCTTGGCGAGCCCGGCTGTACGAGGCTCGCAGCGGCTATCCACGCTTTCAGGCGCTAAAT 1380
 Qy 1431 AAAAGCGCTTTCGACGC 1448
 Db 1381 AAAAGCGCTTTCGACGC 1398

RESULT 5
 ACC45311
 ID ACC45311 standard; DNA; 1278 BP.
 XX
 AC ACC45311;

XX DT 17-JUN-2003 (first entry)
 XX DE Corynebacterium glutamicum eno encoding DNA SEQ ID NO:1.
 XX KW Fine chemical; Coryneform bacteria; Escherichia coli; microorganism;
 KW genetically modified microorganism; metabolite; biosynthesis; amino acid;
 KW vitamin; nucleoside; nucleotide; pigment; protein; human medicine;
 KW pharmaceutical; food; animal feeding; eno; gene; ds.
 XX OS Corynebacterium glutamicum.
 XX FH Location/Qualifiers
 XX CDS 1..1278
 FT /*tag= a
 FT /product= "eno protein"
 XX PN WO2003023016-A2.
 XX PD 20-MAR-2003.
 XX PF 11-SEP-2002; 2002WO-EP010174.
 XX PR 13-SEP-2001; 2001DE-01045043.
 XX PA (DEGS) DEGUSSA AG.
 XX PI Farwick M, Hermann T;
 XX DR WPI; 2003-354534/33.
 XX DR P-PSDB; ABP97001.
 XX DX Microorganism useful for producing e.g. fine chemicals, has permanently
 PT altered phosphorylatability protein, such that biosynthesis of fine
 PT chemical synthesized by microorganism is increased compared to wild-type.
 XX Example 3; Page 78-80; 120pp; English.
 XX The present invention describes a microorganism (I), in which the
 CC phosphorylatability of at least one protein has been permanently altered
 CC such that the biosynthesis of at least one fine chemical synthesised by
 CC the microorganism is increased compared to the wild type. Also described:
 CC (1) use of a DNA (II) sequence coding for a protein which contains a
 CC phosphorylation site, where the sequence contains such a mutation that
 CC the protein is changed in its phosphorylatability for the production of
 CC (1), or for the production of fine chemicals; and (2) a method for
 CC producing fine chemicals or metabolites comprising using (1). (1) is
 CC useful for producing fine chemicals or metabolites, such as amino acids,
 CC vitamins, nucleosides, nucleotides, pigments or proteins. The amino acids
 CC and vitamins produced using (1) can be used in human medicine, in the
 CC pharmaceutical industry, food industry and in animal feeding. (1)
 CC produces larger amount of desired fine chemical or a metabolite than the
 CC wild type. The present sequence encodes wild type eno from
 CC Corynebacterium glutamicum, which is used in an example from the present
 CC invention
 XX SQ Sequence 1278 BP; 271 A; 399 C; 347 G; 261 T; 0 U; 0 Other;

Query Match 80.9%; Score 1276.4; DB 8; Length 1278;
 Best Local Similarity 99.9%; Pred. No. 2.8e-313;
 Matches 1277; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 151 GTGGCTGAATCATGACGATTCGCTCGGAAATTCGACCTCCCGGTAACCCAAAC 210
 Db 1 GTGGCTGAATCATGACGATTCGCTCGGAAATTCGACCTCCCGGTAACCCAAAC 60
 Qy 211 GTCGAGGACAGAGGTTTCTTGGATGACGGTTCACAGGTTCCAGAGTGTTCATCCGGC 270
 Db 61 GTCGAGGACAGAGGTTTCTTGGATGACGGTTCACAGGTTCCAGAGTGTTCATCCGGC 120
 Qy 271 GCATCCACCGGCTCCACAGGCTCATGAGCTCGGTGACGGTGGCGATCGCTACCTGGGC 330
 Db 121 GCATCCACCGGCTCCACAGGCTCATGAGCTCGGTGACGGTGGCGATCGCTACCTGGGC 180

```

Qy 331 AAGGGCGTTTGAAGGCGAGTTGAAACGTCACGAGAAATCGGCGACGAGCTCGCTGGC 390
Db 181 AAGGGCGTTTGAAGGCGAGTTGAAACGTCACGAGAAATCGGCGACGAGCTCGCTGGC 240

Qy 391 CTAGAGGCTGACGATGAGCGCTCATCGAGAGCAATGATCAAGCTTGTATGACCGCC 450
Db 241 CTAGAGGCTGACGATGAGCGCTCATCGAGAGCAATGATCAAGCTTGTATGACCGCC 300

Qy 451 AACAAATCCCGCTGGGTGAAAGCAATCTTGGTGTTCATGCTGCTGTCGAAAGCT 510
Db 301 AACAAATCCCGCTGGGTGAAAGCAATCTTGGTGTTCATGCTGCTGTCGAAAGCT 360

Qy 511 GCTGCTGATTCGCGAGCGCTCCCACTGTTCCGTACATCGGTGACCAACGACACGTT 570
Db 361 GCTGCTGATTCGCGAGCGCTCCCACTGTTCCGTACATCGGTGACCAACGACACGTT 420

Qy 571 CTTCCAGTTCGAATGATGAACATCATCAAGGTGGCGCTCAAGCTGACTCGGTTGAC 630
Db 421 CTTCCAGTTCGAATGATGAACATCATCAAGGTGGCGCTCAAGCTGACTCGGTTGAC 480

Qy 631 GTTCAGGAATTCATGATCGCTCCAAATCGGTGCGAGACCTTCTGAGGCTCTCCGCAAC 690
Db 481 GTTCAGGAATTCATGATCGCTCCAAATCGGTGCGAGACCTTCTGAGGCTCTCCGCAAC 540

Qy 691 GGGCGGAGGTCTACCAAGCACTGAAGTCCGTCTCAAGGAAAGGGCTGTCCACCGGA 750
Db 541 GGGCGGAGGTCTACCAAGCACTGAAGTCCGTCTCAAGGAAAGGGCTGTCCACCGGA 600

Qy 751 CTTGGCGATGAGGCGGCTTTCGCTTCCGTTCGGCTCCACCGGTGAGGCTCTTGACCTT 810
Db 601 CTTGGCGATGAGGCGGCTTTCGCTTCCGTTCGGCTCCACCGGTGAGGCTCTTGACCTT 660

Qy 811 ATCGTTAAGGCAATCGAGAGGCTGGCTTACCCGAGGAGGACATCGCTTGTCTGTG 870
Db 661 ATCGTTAAGGCAATCGAGAGGCTGGCTTACCCGAGGAGGACATCGCTTGTCTGTG 720

Qy 871 GAGCTTCTCTCTGAGTTCCTCAAGGAGGCACTTACCACTTCGAAGTGGCGACAC 930
Db 721 GAGCTTCTCTCTGAGTTCCTCAAGGAGGCACTTACCACTTCGAAGTGGCGACAC 780

Qy 931 TCCGACGCTGAGATGGCAAGCTTTACGCTGAGCTGCTGACGCGTACCAATCGTCTCC 990
Db 781 TCCGACGCTGAGATGGCAAGCTTTACGCTGAGCTGCTGAGGCTACCAATCGTCTCC 840

Qy 991 ATCGAGGACCACTGAGGAGGATGACTGGAGGTTTACCAACCTCACCGCAACATC 1050
Db 841 ATCGAGGACCACTGAGGAGGATGACTGGAGGTTTACCAACCTCACCGCAACATC 900

Qy 1051 GGGGACAAAGTTTCAGATCGTTGGCGAGGCTTCTTCTGTCACCAACCTGAGCGCTGAAG 1110
Db 901 GGGGACAAAGTTTCAGATCGTTGGCGAGGCTTCTTCTGTCACCAACCTGAGCGCTGAAG 960

Qy 1111 GAGGACATCGCTAAGAGGCTGCCAATCTCCATCTCGTTAAGGTGAACAGATCGGTACC 1170
Db 961 GAGGACATCGCTAAGAGGCTGCCAATCTCCATCTCGTTAAGGTGAACAGATCGGTACC 1020

Qy 1171 CTCACGAGACCTTCGACGCTGTCGACATGCTCACCGGAGGCTACCTTCATGATG 1230
Db 1021 CTCACGAGACCTTCGACGCTGTCGACATGCTCACCGGAGGCTACCTTCATGATG 1080

Qy 1231 TCCACCGCTTCGGTGAGACCGGAGGACCAACCACTTCTGACCTCGCAGTTGCACTCAAC 1290
Db 1081 TCCACCGCTTCGGTGAGACCGGAGGACCAACCACTTCTGACCTCGCAGTTGCACTCAAC 1140

Qy 1291 TGTGGCCAGATCAAGACTGTGTCTCCAGACGCTTCCGACCGGTGTCCGAAAGTACAACAG 1350
Db 1141 TGTGGCCAGATCAAGACTGTGTCTCCAGACGCTTCCGACCGGTGTCCGAAAGTACAACAG 1200

Qy 1351 CTTCTCCGCAATCGAGAGCTGCTTGGCGAGCGCGGCTGTACGAGGTGCGAGCGCATTC 1410
Db 1201 CTTCTCCGCAATCGAGAGCTGCTTGGCGAGCGCGGCTGTACGAGGTGCGAGCGCATTC 1260

```

```

Qy 1411 CCACGCTTTTCAGGGCTAA 1428
Db 1261 CCACGCTTTTCAGGGCTAA 1278

```

RESULT 6
AAH66050

ID AAH66050 standard; DNA; 1275 BP.

XX AAH66050;

XX AC (first entry)

XX 26-SEP-2001

XX C glutamicum coding sequence fragment SEQ ID NO: 1085.

XX Corynebacterium; amino acid synthesis; vitamin; saccharide;

XX organic acid synthesis; ds.

XX Corynebacterium glutamicum.

XX BP1108790-A2.

XX 20-JUN-2001.

XX 18-DEC-2000; 2000EP-00127688.

XX 16-DEC-1999; 99JP-00377484.

XX 07-APR-2000; 2000JP-00159162.

XX 03-AUG-2000; 2000JP-00280988.

XX (KYOW) KYOWA HAKKO KOGYO KK.

XX Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;

XX Tateishi N, Senoh A, Ikeda M, Ozaki A;

XX WPI; 2001-376931/40.

XX P-PSDB; AAG90831.

XX Novel polynucleotides derived from Corynebacterium bacteria, for identifying

XX mutation point of a gene, measuring expression of a gene, analyzing

XX expression profile or pattern of a gene and identifying homologous gene.

XX Claim 8; SEQ ID NO 1085; 246pp + Sequence Listing; English.

XX The present invention provides a number of nucleotide and protein

XX sequences from the Corynebacterium bacterium Corynebacterium glutamicum. These

XX are useful for identifying the mutation point of a gene derived from a

XX mutant of corynebacterium bacterium, measuring expression amount and analysing

XX the expression profile or expression pattern of a gene derived from

XX Corynebacterium bacterium, and identifying a homologue of a gene derived from

XX Corynebacterium bacterium. Corynebacterium bacteria are useful for producing amino

XX acids, nucleic acids, vitamins, saccharides and organic acids,

XX particularly L-lysine. The present sequence is a nucleic acid described

XX in the exemplification of the invention. Note: The sequence data for this

XX patent did not form part of the printed specification, but was obtained

XX in electronic format directly from the European Patent Office

XX Sequence 1275 BP; 269 A; 399 C; 347 G; 260 T; 0 U; 0 Other;

XX Query Match 80.7%; Score 1273.4; DB 5; Length 1275;

XX Best Local Similarity 99.9%; Pred. No. 1.7e-312;

XX Matches 1274; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 151 GTGGCTGAATCATGACGATTCGCTCGGAAATTCGACTCCCGCGGTAAACCAACC 210

Db 1 GTGGCTGAATCATGACGATTCGCTCGGAAATTCGACTCCCGCGGTAAACCAACC 60

Qy 211 GTGAGGACAGAGTTTCTCGGATGACGTTCCACGCTGTCGAGGTGTTCATCCGCG 270

Db 61 GTGAGGACAGAGTTTCTCGGATGACGTTCCACGCTGTCGAGGTGTTCATCCGCG 120

Qy 271 GCATCCACCGCGCTCCACGAGGCTCATGAGCTGCGTGCAGGTGGCGATCGCTACCTGGCG 330

Db 121 GCATCCACGGCGCTCACGAGGCTCATGAGTCGGTGACGGTGGCGATCGCTACCTGGGC 180
Qy 331 AAGGCGCTTTGAAGGCGAGTTGAAAAAGTCAACGAAGAAATCGCGCAGAGCTCGTGGC 390
Db 181 AAGGCGCTTTGAAGGCGAGTTGAAAAAGTCAACGAAGAAATCGCGCAGAGCTCGTGGC 240
Qy 391 CTAGAGGCTCAGCATCAGCGCTCATCAGCGAGCAATGATCAAGCTTGATGCGACCGCC 450
Db 241 CTAGAGGCTCAGCATCAGCGCTCATCAGCGAGCAATGATCAAGCTTGATGCGACCGCC 300
Qy 451 AACAACTCCGCGCTGGGTCAAAAGCAATCTTGGTGTTCATGCTGTTCGAAAGGCT 510
Db 301 AACAACTCCGCGCTGGGTCAAAAGCAATCTTGGTGTTCATGCTGTTCGAAAGGCT 360
Qy 511 GCTGCTGATTCGCGAGGCTCCCACTGTTCCGCTACATCGGTGGAACAAAGCACAGTT 570
Db 361 GCTGCTGATTCGCGAGGCTCCCACTGTTCCGCTACATCGGTGGAACAAAGCACAGTT 420
Qy 571 CTTCCAGTTCCAAATGATGAACATCATCAACGGTGGGCTCACGCTGACTCCGGTGTGAC 630
Db 421 CTTCCAGTTCCAAATGATGAACATCATCAACGGTGGGCTCACGCTGACTCCGGTGTGAC 480
Qy 631 GTTCAGGAATTCATGATCGCTCCAATCGGTGCAGAGACCTTCTCTGAGGCTCTCCGCAAC 690
Db 481 GTTCAGGAATTCATGATCGCTCCAATCGGTGCAGAGACCTTCTCTGAGGCTCTCCGCAAC 540
Qy 691 GGGCGGAGGCTTACACGCACTGAAGTCCGCTGATCAAGGAAAGGGCCCTGTCCACGGA 750
Db 541 GGGCGGAGGCTTACACGCACTGAAGTCCGCTGATCAAGGAAAGGGCCCTGTCCACGGA 600
Qy 751 CTTGGCGATCAGGCGGCTTCGCTCCCTTCGCTGGCTCCACCGTGAGGCTCTTGACCTT 810
Db 601 CTTGGCGATCAGGCGGCTTCGCTCCCTTCGCTGGCTCCACCGTGAGGCTCTTGACCTT 660
Qy 811 ATCGTTAAGCAATCGAAGAGGCTGGCTTACCCCGAGGCAAGGACATCGCTCTTGCTCTG 870
Db 661 ATCGTTAAGCAATCGAAGAGGCTGGCTTACCCCGAGGCAAGGACATCGCTCTTGCTCTG 720
Qy 871 GAGTTGCTTCCTGAGTTCTTCAAGAGCGCACCTTACACTTCGAGGTGGCGAGCAC 930
Db 721 GAGTTGCTTCCTGAGTTCTTCAAGAGCGCACCTTACACTTCGAGGTGGCGAGCAC 780
Qy 931 TCCGAGCTGAGATGGCAACGTTTACGCTGAGCTCGTTGACGCGTACCCAATCGTCTCC 990
Db 781 TCCGAGCTGAGATGGCAACGTTTACGCTGAGCTCGTTGACGCGTACCCAATCGTCTCC 840
Qy 991 ATCGAGGACCACTGCAGGAAGATGACTGGAGGGTTTACACCACTTACCGCAACCCATC 1050
Db 841 ATCGAGGACCACTGCAGGAAGATGACTGGAGGGTTTACACCACTTACCGCAACCCATC 900
Qy 1051 GGGCACAAGGTTGAGTCTGTTGGGAGGACTTCTGTCACCAACCTGAGCGGCTGAAG 1110
Db 901 GGGCACAAGGTTGAGTCTGTTGGGAGGACTTCTGTCACCAACCTGAGCGGCTGAAG 960
Qy 1111 GAGGCACTCGCTAAGAGGCTGCCAATCCATCTGTTTAAAGTGAACCAAGATCGGTACC 1170
Db 961 GAGGCACTCGCTAAGAGGCTGCCAATCCATCTGTTTAAAGTGAACCAAGATCGGTACC 1020
Qy 1171 CTCACGAGACCTTCGACGCTGTCGACATGGCTCACCGCGAGGCTACACCTCCATGATG 1230
Db 1021 CTCACGAGACCTTCGACGCTGTCGACATGGCTCACCGCGAGGCTACACCTCCATGATG 1080
Qy 1231 TCCACGCTTCGGTGAAGCGGAGGACACCACTTCTGCTGACCTCGAGTTGCACTCAAC 1290
Db 1081 TCCACGCTTCGGTGAAGCGGAGGACACCACTTCTGCTGACCTCGAGTTGCACTCAAC 1140
Qy 1291 TGTGGCCAGATCAAGACTGGTGTCCAGCAGCTTCCGACCGTGTCCGAAAGTACCAAC 1350
Db 1141 TGTGGCCAGATCAAGACTGGTGTCCAGCAGCTTCCGACCGTGTCCGAAAGTACCAAC 1200
Qy 1351 CTTCTCGGATCGAGCAGCTGCTTGGCGAGCGCGGCGTCTACGAGGTTCGACGGCATTC 1410
Db 1201 CTTCTCGGATCGAGCAGCTGCTTGGCGAGCGCGGCGTCTACGAGGTTCGACGGCATTC 1260

Qy 1411 CCACGCTTTCAGGGC 1425
Db 1261 CCACGCTTTCAGGGC 1275

RESULT 7

ACC45312
ID ACC45312 standard; DNA; 1299 BP.

XX ACC45312;

DT 17-JUN-2003 (first entry)

XX Corynebacterium glutamicum eno DNA fragment SEQ ID NO:3.

XX Fine chemical; Coryneform bacteria; Escherichia coli; microorganism;
KW genetically modified microorganism; metabolite; biosynthesis; amino acid;
KW vitamin; nucleoside; nucleotide; pigment; protein; human medicine;
KW pharmaceutical; food; animal feeding; eno; gene; ds.

XX Corynebacterium glutamicum.

OS

XX WO2003023016-A2.

PN

XX 20-MAR-2003.

XX 11-SEP-2002; 2002WO-EP010174.

PF

XX 13-SEP-2001; 2001DE-01045043.

PR

XX (DEGS) DEGUSSA AG.

XX Farwick M, Hermann T;

XX WPI; 2003-354534/33.

XX Microorganism useful for producing e.g. fine chemicals, has permanently

PT altered phosphorylatability protein, such that biosynthesis of fine

PT chemical synthesized by microorganism is increased compared to wild-type.

XX Example 3; Page 81-82; 120pp; English.

XX The present invention describes a microorganism (I), in which the

XX phosphorylatability of at least one protein has been permanently altered

XX such that the biosynthesis of at least one fine chemical synthesised by

XX the microorganism is increased compared to the wild type. Also described:

XX (1) use of a DNA (II) sequence coding for a protein which contains a

XX phosphorylation site, where the sequence contains such a mutation that

XX the protein is changed in its phosphorylatability for the production of

XX (I), or for the production of fine chemicals; and (2) a method for

XX producing fine chemicals or metabolites comprising using (I). (I) is

XX useful for producing fine chemicals or metabolites, such as amino acids,

XX vitamins, nucleosides, nucleotides, pigments or proteins. The amino acids

XX and vitamins produced using (I) can be used in human medicine. (I)

XX pharmaceutical industry, food industry and in animal feeding. (I)

XX produces larger amount of desired fine chemical or a metabolite than the

XX wild type. The present sequence represents a wild type eno fragment DNA

XX sequence from Corynebacterium glutamicum, which is used in an example

XX from the present invention

XX SQ Sequence 1299 BP; 274 A; 409 C; 353 G; 263 T; 0 U; 0 Other;

Query Match 80.7%; Score 1273.4; DB 8; Length 1299;
Best Local Similarity 99.9%; Pred. No. 1.7e-312;
Matches 1274; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 154 GCTGAATCATGCACGATTCGCTCCGAAATTCGACTCCCGGTAACCAACGTC 213

Db 10 GCTGAATCATGCACGATTCGCTCCGAAATTCGACTCCCGGTAACCAACGTC 69

Qy 214 GAGGCAGAGGTTTCTTGGATGACGTTCCCGGTTCCCGGTTTCCATCCCGGCA 273

Query Match	80.6%	Score 1271.6	DB 8	Length 1278
Best Local Similarity	99.7%	Pred. No. 4.8e-312		
Matches 1274; Conservative 0;	Mismatches 4;	Indels 0;	Gaps 0;	
QY	151	GTGGCTCAAAATCATGACAGTATTTCGCCTCGGAAAATTTCCGACTCCGCGGTAAACCCAAC	210	
Db	1	GTGGCTGAATATCATGACAGTATTTCGCCTCGGAAAATTTCCGACTCCGCGGTAAACCCAAC	60	
QY	211	GTGAGGCAGAGGTTTTCTTGATGACGGTTTTCCCACGGTGTCCGAGGTGTTCCTATCCGGC	270	
Db	61	GTGAGGCAGAGGTTTTCTTGATGACGGTTTTCCCACGGTGTCCGAGGTGTTCCTATCCGGC	120	
QY	271	GCATCCACCGGCTCCACGAGGCTCATGAGCTGCGTAGCGTACGTCGGCATCGTACTCGGCG	330	
Db	121	GCATCCACCGGCTCCACGAGGCTCATGAGCTGCGTAGCGTACGTCGGCATCGTACTCGGCG	180	
QY	331	AAGGGCGTTTTTAAGGCAGTTGAAAAAGTCAAAGAAATTCGGCGACGAGCTCGCTGGC	390	
Db	181	AAGGGCGTTTTTAAGGCAGTTGAAAAAGTCAAAGAAATTCGGCGACGAGCTCGCTGGC	240	
QY	391	CTAGAGCTCAGCATGAGCGCTCATCGAAGAACAATGATCAAGCTTGATGGCACCGCC	450	
Db	241	CTAGAGCTCAGCATGAGCGCTCATCGAAGAACAATGATCAAGCTTGATGGCACCGCC	300	
QY	451	AACAAGTCCCGCTGGGTGCAACCGCAATCCTTCGGTGTTCATAGGCTGTGTGCAAGGCT	510	
Db	301	AACAAGTCCCGCTGGGTGCAACCGCAATCCTTCGGTGTTCATAGGCTGTGTGCAAGGCT	360	
QY	511	GTGCTGATTCGCGAGGCCCTCCCACTGTTTCGCTACATCGGTGACCAACGCAACGCTT	570	
Db	361	GCTGCTGATTCGCGAGGCCCTCCCACTGTTTCGCTACATCGGTGACCAACGCAACGCTT	420	
QY	571	CTTCCAGTTCCAATGATGAACATCATCAACGGTGGCGCTCAGCTGACTCGGTTGTGAC	630	
Db	421	CTTCCAGTTCCAATGATGAACATCATCAACGGTGGCGCTCAGCTGACTCGGTTGTGAC	480	
QY	631	GTTTCAGGAATTCATGATCGCTCCAATCGGTGCAGAGACCTTCTCTGAGGCTCTCCGCAAC	690	
Db	481	GTTTCAGGAATTCATGATCGCTCCAATCGGTGCAGAGACCTTCTCTGAGGCTCTCCGCAAC	540	
QY	691	GGCGCGAGGTCTACACGCACTGAAGTTCGTCATCAAGGAAAGGCGCTGTTCACCGGA	750	
Db	541	GGCGCGAGGTCTACACGCACTGAAGTTCGTCATCAAGGAAAGGCGCTGTTCACCGGA	600	
QY	751	CTTGGCGATCAGGCGGCTTCGCTCTTCGTCGGCTCCACCCTGAGGCTCTTGACCTT	810	
Db	601	CTTGGCGATCAGGCGGCTTCGCTCTTCGTCGGCTCCACCCTGAGGCTCTTGACCTT	660	
QY	811	ATCGTTAAGGCAATCGAAGAGGTGGCTTCACCCCAGGCAAGCAATCGCTCTTGCTCTG	870	
Db	661	ATCGTTAAGGCAATCGAAGAGGTGGCTTCACCCCAGGCAAGCAATCGCTCTTGCTCTG	720	
QY	871	GAGCTTGCTTCTCTGAGTCTTCAAGGACGGCACTTACCACTTCGAAGTGGCGACAC	930	
Db	721	GAGCTTGCTTCTCTGAGTCTTCAAGGACGGCACTTACCACTTCGAAGTGGCGACAC	780	
QY	931	TCCGCACTGAGATGGCAAAAGTGTTCACGCTGAGCTCGTTGACGGTACCCAATCGTCTCC	990	
Db	781	TCCGCACTGAGATGGCAAAAGTGTTCACGCTGAGCTCGTTGACGGTACCCAATCGTCTCC	840	
QY	991	ATCGAGGACCACTGCAGGAGATGACTGGAGGGTTACCAACCTTACCGCAACCATC	1050	
Db	841	ATCGAGGACCACTGCAGGAGATGACTGGAGGGTTACCAACCTTACCGCAACCATC	900	
QY	1051	GGCGCAAGGTTTCAGATCGTTGGCGACGACTTCTTCGTCACCAACCTGAGCGGCTGAAG	1110	
Db	901	GGCGCAAGGTTTCAGATCGTTGGCGACGACTTCTTCGTCACCAACCTGAGCGGCTGAAG	960	
QY	1111	GAGGGCATCTCTAAGAGGCTGCCAATCCATCTCTGTTTAAGTGAAACCAATCGGTACC	1170	
Db	961	GAGGGCATCTCTAAGAGGCTGCCAAGAGATCTCTGTTTAAGTGAAACCAATCGGTACC	1020	

QY	1171	CTCACCGAGAC	CTTCGACGCTGTGCGACATGGCTCACCGCGAGGCTACACCTTCATGATG	12350
Db	1021	CTCACCGAGAC	CTTCGACGCTGTGCGACATGGCTCACCGCGAGGCTACACCTTCATGATG	1080
QY	1231	TCCACACGCTTCCGGTGTGAGACCGAGGACACACCACTTCTGACCTCGCAGTTGCACTCAAC	1290	
Db	1081	TCCACACGCTTCCGGTGTGAGACCGAGGACACCACTTCTGACCTCGCAGTTGCACTCAAC	1140	
QY	1291	TGTGGCCAGATCAAGACTGTGTCTCCAGCAGCTTCCGACCGTGTGCGAAAGTACAACCCAG	1350	
Db	1141	TGTGGCCAGATCAAGACTGTGTCTCCAGCAGCTTCCGACCGTGTGCGAAAGTACAACCCAG	1200	
QY	1351	CTTCTCCGATCGAGCAGCTGCTTGGCGAGCGCGCGCTCTACGAGGTGCGAGGGCATTC	1410	
Db	1201	CTTCTCCGATCGAGCAGCTGCTTGGCGAGCGCGCGCTCTACGAGGTGCGAGGGCATTC	1260	
QY	1411	CCACGCTTTTCAGGGCTAA	1428	
Db	1261	CCACGCTTTTCAGGGCTAA	1278	
RESULT 9				
AC	ACC45313	ID ACC45313 standard; DNA; 1299 BP.		
XX	AC	ACC45313;		
XX	AC	ACC45313;		
DT	17-JUN-2003	(first entry)		
XX	Modified	eno DNA fragment S330E SEQ ID NO:4.		
DE	Fine chemical;	Coryneform bacteria; Escherichia coli; microorganism;		
KW	Genetically modified microorganism;	metabolite; biosynthesis; amino acid;		
KW	vitamin; nucleoside; nucleotide; pigment; protein; human medicine;	pharmaceutical; food; animal feeding; eno; gene; ds.		
KW	pharmaceutical; food; animal feeding; eno; gene; ds.	Corynebacterium glutamicum.		
XX	Corynebacterium glutamicum.	Synthetic.		
OS	Synthetic.	WO2003023016-A2.		
XX	PN	20-MAR-2003.		
XX	PD	11-SEP-2002; 2002WO-EP010174.		
XX	PF	13-SEP-2001; 2001DE-01045043.		
XX	PR	(DEGS) DEGUSSA AG.		
PA	Farwick M, Hermann T;	WPI; 2003-354534/33.		
XX	PI	Microorganism useful for producing e.g. fine chemicals, has permanently		
XX	DR	altered phosphorylatability protein, such that biosynthesis of fine		
XX	PT	chemical synthesized by microorganism is increased compared to wild-type.		
XX	PS	Claim 17; Page 92; 120pp; English.		
XX	CC	The present invention describes a microorganism (I), in which the		
CC	CC	phosphorylatability of at least one protein has been permanently altered		
CC	CC	such that the biosynthesis of at least one fine chemical synthesised by		
CC	CC	the microorganism is increased compared to the wild type. Also described:		
CC	CC	(1) use of a DNA (II) sequence coding for a protein which contains a		
CC	CC	phosphorylation site, where the sequence contains such a mutation that		
CC	CC	the protein is changed in its phosphorylatability for the production of		
CC	CC	(I), or for the production of fine chemicals; and (2) a method for		
CC	CC	producing fine chemicals or metabolites comprising using (I). (I) is		
CC	CC	useful for producing fine chemicals or metabolites, such as amino acids,		
CC	CC	vitamins, nucleosides, nucleotides, pigments or proteins. The amino acids		
CC	CC	and vitamins produced using (I) can be used in human medicine, in the		
CC	CC	pharmaceutical industry, food industry and in animal feeding. (I)		
CC	CC	produces larger amount of desired fine chemical or a metabolite than the		

CC wild type. The present sequence represents a modified eno fragment DNA
 CC sequence from Corynebacterium glutamicum, which is used in an example
 CC from the present invention
 XX
 SQ Sequence 1299 BP; 275 A; 407 C; 355 G; 262 T; 0 U; 0 Other;

Query Match 80.4%; Score 1268.6; DB 8; Length 1299;
 Best Local Similarity 99.7%; Pred. No. 2.8e-311;
 Matches 1271; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 154 GCTGAATCATGACGATTTCCGCGAATTTCTCGACTCCCGCGGTAAACCCACCGTC 213
 Db 10 GCTGAATCATGACGATTTCCGCGAATTTCTCGACTCCCGCGGTAAACCCACCGTC 69

Qy 214 GAGCAGAGGTTTCCGTGGATGACGGTTCCACGGTGTGCGAGGTGTTTCCATCCGGGCA 273
 Db 70 GAGCAGAGGTTTCCGTGGATGACGGTTCCACGGTGTGCGAGGTGTTTCCATCCGGGCA 129

Qy 274 TCCACCGCGTCCACGAGGCTCATGAGTGGTGAACGGTGGCGATCGCTACCTGGGCAAG 333
 Db 130 TCCACCGCGTCCACGAGGCTCATGAGTGGTGAACGGTGGCGATCGCTACCTGGGCAAG 189

Qy 334 GGGCTTTTGAAGGCGAGTTGAAACGTCACGAAAGATTCGGCGAGCGCTCGTGGGCTA 393
 Db 190 GGGCTTTTGAAGGCGAGTTGAAACGTCACGAAAGATTCGGCGAGCGCTCGTGGGCTA 249

Qy 394 GAGGCTGACGATCAGCGCTCATCGAGGAAGCAATGATCAAGCTTGATGGCACCGCAAC 453
 Db 250 GAGGCTGACGATCAGCGCTCATCGAGGAAGCAATGATCAAGCTTGATGGCACCGCAAC 309

Qy 454 AAGTCCCGCTGGTGCAGGCAATCTTGGTGTTCATGGCTGTGCAAAAGGCTGCT 513
 Db 310 AAGTCCCGCTGGTGCAGGCAATCTTGGTGTTCATGGCTGTGCAAAAGGCTGCT 369

Qy 514 GCTGATTCGCGAGCGCTCCCACTGCTCGGTACATCGGTGACGACCAAGCAGCATGTTCT 573
 Db 370 GCTGATTCGCGAGCGCTCCCACTGCTCGGTACATCGGTGACGACCAAGCAGCATGTTCT 429

Qy 574 CCAGTTCCAATGATGAACATCATCAACGGTGGCGCTCAGCGTGAATCCGGTGTGACGTT 633
 Db 430 CCAGTTCCAATGATGAACATCATCAACGGTGGCGCTCAGCGTGAATCCGGTGTGACGTT 489

Qy 634 CAGGAATTCATGATCGCTCCCAATCGGTGCAGAGACCTTCTCTGAGGCTCTCCGCAACGGC 693
 Db 490 CAGGAATTCATGATCGCTCCCAATCGGTGCAGAGACCTTCTCTGAGGCTCTCCGCAACGGC 549

Qy 694 GCGAGGTCTACACGACCTGAAGTCCGTATCAAGAAAGGCGCTGTCCACCGACTT 753
 Db 550 GCGAGGTCTACACGACCTGAAGTCCGTATCAAGAAAGGCGCTGTCCACCGACTT 609

Qy 754 GCGATGAGGCGGCTTCCGCTTCCGTCGGCTCCACCGTGAAGGCTCTTGACCTTATC 813
 Db 610 GCGATGAGGCGGCTTCCGCTTCCGTCGGCTCCACCGTGAAGGCTCTTGACCTTATC 669

Qy 814 GTTAAGCAATTCGAGAGGCTGGCTTCAACCGGCAAGGACATCGCTTGTCTCTGGAC 873
 Db 670 GTTAAGCAATTCGAGAGGCTGGCTTCAACCGGCAAGGACATCGCTTGTCTCTGGAC 729

Qy 874 GTTGCTTCTCTGAGTTCCTCAAGGACGGACCTTACCACTTCGAAAGTGGCGACATCC 933
 Db 730 GTTGCTTCTCTGAGTTCCTCAAGGACGGACCTTACCACTTCGAAAGTGGCGACATCC 789

Qy 934 GCAGCTGAGATGCAACGTTTACGCTGAGCTGTTGACCGGTACCAATCGTCTCCATC 993
 Db 790 GCAGCTGAGATGCAACGTTTACGCTGAGCTGTTGACCGGTACCAATCGTCTCCATC 849

Qy 994 GAGGACCACTGCGAGGAAGTACTGGGAGGGTTTACACCAACCTCAGCGCAACCATCGG 1053
 Db 850 GAGGACCACTGCGAGGAAGTACTGGGAGGGTTTACACCAACCTCAGCGCAACCATCGG 909

Qy 1054 GACAAGTTTCAGATCGTTGGCGACGATTTCTGCTACCAACCTTGAGCGCTTGAAGGAG 1113
 Db 910 GACAAGTTTCAGATCGTTGGCGACGATTTCTGCTACCAACCTTGAGCGCTTGAAGGAG 969

Qy 1114 GGATCGCTTAAGAGGCTGCCAATCTCCATCCTGTTAAGGTGAACACGATCGGTACCCCTC 1173
 Db 970 GGATCGCTTAAGAGGCTGCCAATCTCCATCCTGTTAAGGTGAACACGATCGGTACCCCTC 1029

Qy 1174 ACCGAGACCTTCGACGCTGCGACATGGCTCACCGCGAGGCTACACCTTCATGATGCC 1233
 Db 1030 ACCGAGACCTTCGACGCTGCGACATGGCTCACCGCGAGGCTACACCTTCATGATGCC 1089

Qy 1234 CACGTTCCGCTGAGACCGAGGACACCACTTCTGACCTTCGCGAGTTGCACTCAACTGT 1293
 Db 1090 CACGTTCCGCTGAGACCGAGGACACCACTTCTGACCTTCGCGAGTTGCACTCAACTGT 1149

Qy 1294 GGCAGATCAAGACTGGTGTCTCCAGCACGTTTCCGACCGTGTCCGAAAGTACAACGAGCTT 1353
 Db 1150 GGCAGATCAAGACTGGTGTCTCCAGCACGTTTCCGACCGTGTCCGAAAGTACAACGAGCTT 1209

Qy 1354 CTCCGATCGAGCAGCTGCTTGGCGACCGCGGCTGTACGAGGTGCGAGGCGATTTCCA 1413
 Db 1210 CTCCGATCGAGCAGCTGCTTGGCGACCGCGGCTGTACGAGGTGCGAGGCGATTTCCA 1269

Qy 1414 CGCTTTCAGGGCTAA 1428
 Db 1270 CGCTTTCAGGGCTAA 1284

RESULT 10
 ACA29854
 ID ACA29854 standard; DNA; 1275 BP.
 XX ACA29854;
 AC AC
 XX XX
 DT 19-JUN-2003 (first entry)
 XX
 DE Prokaryotic essential gene #11511.
 XX
 KW Antisense; ds; prokaryotic essential gene; cell proliferation;
 KW drug design; gene.
 XX
 OS Corynebacterium diphtheriae.
 XX
 PN WO200277183-A2.
 XX
 PD 03-OCT-2002.
 XX
 PF 21-MAR-2002; 2002WO-US009107.
 XX
 PR 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 08-MAR-2002; 2002US-0362699P.
 XX
 PA (ELIT-) ELITRA PHARM INC.
 XX
 PI Wang L, Zamudio C, Malone C, Hasebeck R, Ohlsen KL, Zyskind JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX
 DR WPI; 2003-029926/02.
 DR P-PSDB; ABU25984.
 XX
 PT New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX
 PS Claim 14; SEQ ID NO 17724; 1766bp; English.
 CC
 CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated

CC polypeptide or its fragment whose expression is inhibited by the
CC antiseptic nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antiseptic nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
CC prokaryotic essential genes. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 1275 BP; 268 A; 378 C; 354 G; 275 T; 0 U; 0 Other;

Query Match 57.5%; Score 907; DB 8; Length 1275;
Best Local Similarity 82.0%; Pred. No. 1.4e-219;
Matches 1045; Conservative 0; Mismatches 230; Indels 0; Gaps 0;

QY 151 GTGGCTGAATCATGACGATTCGCTCGCGAAATTCCTGACTCCCGCGGTAAACCCAAAC 210
DB 1 GTGGCTGACATTAATGACGATTTTGTCTCGTGAATCTTGGACTCCCGCGGTAAACCCAAAC 60
QY 211 GTGAGGCTGACGATTTTCTCGGATGACGTTTCCACGGTGTCCGAGTGTTCATCCGGC 270
DB 61 GTTGAAGCCGAAGTTTTTCTTGGATGACGATCCACGGCGTTTGGGGCGTTCTTCCGGT 120
QY 271 GCATCCACCGCGGTCCACGAGGCTCATGAGTGTGCGTGACGGTGGCGATCGCTACCTGGGC 330
DB 121 GCGTCCACCGGTTTACGAGGCTCACGAGCTTCGCGACGGTGGCGAGCGCTACCTCGGC 180
QY 331 AAGGGCGTTTGAAGGCGATTGAAACGTCACGAAGAAATCGCGACGAGCTCGCTGGC 390
DB 181 AAGGGCGTTTCAATGAGTTAAACGTCGAACGAGGAAATCGCTGACGCAATCGCTGGT 240
QY 391 CTAGAGGCTGACGATCAGCGCTCATCGAAGCAATGATCAAGCTTGTATGGCACCGCC 450
DB 241 GCAGAAGCCGACGATCAGCGCTGATCGATCAGGCGATGATGCTCTCGACGGCATTGAG 300
QY 451 AACAAAGTCCCGCTGGGTGCAAGCGCAATCTTGGTGTTCATGGCTGTGTGCAAGGCT 510
DB 301 AACAAAGTCTCGTCTCGCGCAAGCGCTATCTTGGGTGTATCCATCGCGGTAGCTAAGGCT 360
QY 511 GCTGCTGATTCGAGCGCTCCCACTGTTCCGCTACATCGGTGACCAAAACGACAGCTT 570
DB 361 GCGCGAGAGTCTGTCGGCTACCTTTGTACCGCTACATCGCGCGCCCTAAACGCTCAAGTT 420
QY 571 CTTCAGTTCCATGATGAACATCATCAAGGTGGCGCTCACGCTGACTCCGGTGTGTGAC 630
DB 421 CTTCAGTTCTATGATGAACATTTTAAAGGTGGCGCACACGCTGACTCCGGGTGTGAT 480
QY 631 GTTCAGGAATTCATGATCGCTTCCAATCGGTGCAGAGACCTTCTCTGAGGCTCTCCGCAAC 690
DB 481 GTTCAGGAGTTTATGATGCTCTTATCGGTGCCGAGTCTTCTCTGAGGCTCTGCGCATG 540
QY 691 GCGCGAGGTTCTACCAAGCACTGAAGTCCGTCATCAAGGAAGAGGCGCTGTCCACCGGA 750
DB 541 GGTGCAGAGGTTCTACCACTCTCTTGAAGTCCGTCATTAAGTCCAAAGGAGCTTTCCACCGGC 600
QY 751 CTTGGCGATGAGGCGCGCTTGGCTCTCTTCCGTCGGCTCCACCGCTGAGGCTCTTTCGACCTT 810

DB 601 CTCGGCGACGAAGTGTGTTTCGCACTTCTGTGTAGTCCACCAAGCGAGCTCTCGACCTC 660
QY 811 ATCGTTAAAGCAATCGAAGAGGCTGGCTTCAACCCAGGCAAGGACATTCGCTTGTGCTGTG 870
DB 661 ATCGTTGAGGCAATTCGAAGAGGCTGGCTTCAAGCCAGGCTGCTGACATTCGCACTTGCAC 720
QY 871 GAGGTTGCTTCTCTGAGTCTTCAAGGAGGCGACCTTACCACTTCGCAAGGTGCGCAGCAC 930
DB 721 GAGGTTGCTTCTCTGAGTCTTCAAGGAGTGGCAAGTACCACTTCGAGGAGGCGGAGCAC 780
QY 931 TCCGACAGCTGAGATGTCGCAACCGTTTACGCTGAGCTGTTGACGCGTACCCCAATTCGCTCC 990
DB 781 ACGCTGAGGAGATGCAAGGCTTACGAGCAGCTCATCGTGAGTACCAATTTGTTTCC 840
QY 991 ATCGAGGACCCACTGCGAGGAAGATGATCGGAGGGTTTACCAACCTCACCGCAACATC 1050
DB 841 ATCGAGGACCCACTGCGAGGAAGACGACTGGAGGGCTACACCGCCCTGACCGCGCAATC 900
QY 1051 GCGCAAGGTTTCAAGTCTTGGCGACGACTTCTTCTGTCACCAACCGCTGAGCGCTGAAG 1110
DB 901 GGTGCAAGGTTTCAAGTCTTGGCGACGACTTCTTCTGTCACCAACCGCAGCACCGCTCAAG 960
QY 1111 GAGGCGATCGCTTAAAGAGGCTGCCAACTCCATCTCTGTTAAAGGTGAACAGATCGGTACC 1170
DB 961 GAAGGATCGAGAGAGGCTGCCAATGCTTGTGTTAAAGGTCAACAGATCGGTACC 1020
QY 1171 CTCACCGAGACTTCCGACGCTGTCGATGGCTCACCGCGAGGCTTACACCTCCATGATG 1230
DB 1021 CTGACCGAGACTTTCGACGCTGTTGATCTCGCACACCGCAACCGCTTACCGCACCATGATG 1080
QY 1231 TCCACCGGTTCCGGTGGAGCCGAGGACACCACTTGTGCTGACCTCGCAGTTGCACCTCAAC 1290
DB 1081 TCCACCGGTTCCGGCGAGACTGAAGACACCACTTGTGCTGACCTTGCAGTTCGATTTGGGC 1140
QY 1291 TGTGGCAGATCAAGACTGCTGCTCCAGCACGTTTCCGACCGTGTCCGAAAGTACAACACAG 1350
DB 1141 TGTGGCAGATCAAGACCGGTGCACAGCTCGTTCCGAGCGGTTGCCAAGTACAATCAG 1200
QY 1351 CTTCTCCGATCGAGCAGCTGCTTGGCGACGCGCGCTCTACGCAAGTTCGACGCGATTC 1410
DB 1201 CTTCTCCGATCGAGCAGCAGCTTGCATGACGAGTCTTACGCGAGGCGGTTTCAGCATTT 1260
QY 1411 CCACGGCTTTCAGGCG 1425
DB 1261 CCACGGTTTTCAGGCG 1275

RESULT 11

ADS56353

ID ADS56353 standard; cDNA; 1269 BP.

XX AC ADS56353;

XX AC ADS56353;

DT 02-DEC-2004 (first entry)

XX Bacterial polynucleotide #8340.

DE Bacterial polynucleotide #8340.

XX Bacteria.

XX Bacteria.

KW Recombinant DNA construct; transformed plant; improved plant property;

KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;

KW pathogen tolerance; pest tolerance; plant disease resistance;

KW cell cycle pathway modification; plant growth regulator;

KW homologous recombination; seed oil yield; protein yield; carbohydrate;

KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;

XX bacterial polynucleotide; gene; ss.

OS Bacteria.

XX Bacteria.

XX Bacteria.

PN US2003233675-A1.

XX 18-DEC-2003.

XX 18-DEC-2003.

XX 20-FEB-2003; 2003US-00369493.

XX 21-FEB-2002; 2002US-0360039P.
 XX (CAOY/) CAO Y.
 PA (HINK/) HINKLE G J.
 PA (SLAT/) SLATER S C.
 PA (CHEN/) CHEN X.
 PA (GOLD/) GOLDMAN B S.
 XX
 PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
 XX WPI; 2004-061375/06.
 XX
 PT New recombinant DNA construct comprising a promoter positioned to provide
 PT for expression of a polynucleotide encoding a polypeptide from a
 PT microbial source, useful for producing plants with improved properties.
 XX
 PS Claim 1; SEQ ID NO 32027; 122pp; English.
 XX
 CC The invention relates to a recombinant DNA construct comprising a
 CC promoter functional in a plant cell, where the promoter is positioned to
 CC provide for expression of a polynucleotide encoding a polypeptide from a
 CC microbial source. The invention also relates to a transformed plant
 CC comprising the recombinant DNA construct and a method of producing a
 CC transformed plant having an improved property. The plant is a crop plant
 CC such as maize or soybean. The method of producing a transformed plant
 CC having an improved property comprises transforming a plant with the
 CC recombinant DNA construct and growing the transformed plant, where the
 CC polynucleotide or polypeptide is useful for improving plant properties.
 CC The recombinant DNA construct is useful for producing plants with
 CC improved plant properties, e.g. improved cold, heat or drought tolerance,
 CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
 CC increased resistance to plant disease, better growth rate by modification
 CC of the cell cycle pathway with plant growth regulators, increased rate of
 CC homologous recombination, modified seed oil or protein yield and/or
 CC content, improved yield by modification of carbohydrate, nitrogen or by
 CC phosphorus use and/or uptake, by modification of photosynthesis or by
 CC providing improved plant growth and development under at least one stress
 CC condition, improved lignin production or improved galactomannan
 CC production. This sequence represents a bacterial polynucleotide used in
 CC the scope of the invention. Note: The sequence data for this patent did
 CC not form part of the printed specification but was obtained in electronic
 CC format from USPTO at seqdata.uspto.gov/sequence.html.
 XX
 SQ Sequence 1269 BP; 227 A; 421 C; 443 G; 178 T; 0 U; 0 Other;
 Query Match 43.0%; Score 678.2; DB 13; Length 1269;
 Best Local Similarity 71.0%; Pred. No. 1.4e-161;
 Matches 899; Conservative 0; Mismatches 368; Indels 0; Gaps 0;
 QY 152 TGCGTGAATCATGCGAGTATTCGCTCGCGAATTCCTCGACTCCCGGGTAAACCCACCG 211
 DB 2 TGCGGTGATCGAGGCGAGTACGCGCGGAGATTCGACTCCCGGGTAAACCCACCG 61
 QY 212 TCAGGCGAGAGGTTTCTCGATGACGCGTCCACGCGTTCGAGGTGTTCCATCCGCGC 271
 DB 62 TCAGGTTCGAGTTCGACTCGACGCGGAGGATTCGCGGGGCGGTGTCCTCCAGTGGC 121
 QY 272 CATCCACCGCGCTCCAGAGGCTCATGAGTTCGCTGACGCTGCGGATTCGCTACCTGGGCA 331
 DB 122 CGTCCACCGGTCGATTCGAAGCGGTTCGAACTGCGCGATGCTGGCGACCGTTACGGCGGTA 181
 QY 332 AGGCGGTTTGAAGGCGTTGAAACGTCACAGGAAATCGGGGCGAGCTCGTGGCC 391
 DB 182 AGGCGTGGAGAGGCGGTCCCGCGGTTCACAGGAAATCTCCGAACAGATCGTGGGTT 241
 QY 392 TAGAGGCTGACGATCAGCGCTCATCGACGAGCAATGATGATGAGTGGCACCGCCA 451
 DB 242 ACAGGCGGAGGAGCGGCTGTCGACAGCGCCCTGATCACTCGACGCTGATCCCGC 301
 QY 452 ACAAGTCCCGCTCGGCTGCAACGCAATCCTTGGTGTTCATGGCTGTTCGAAAGGCTG 511
 DB 302 ACAAGTCCCGCTCGGCGGCAACGCAATCCTTGGTGTTCATGGCTGTTCGAAAGGCTG 511

QY 512 CTGCTGATTCGCGAGGCTCCCACTGTTCCGCTACATCGGTGGACCAACGACACGTTTC 571
 DB 362 CCGCGGAGAGCGGAGACTCGCGCTCTTCGCTACCTGGGCGGCCGAAACCCACGTCGC 421
 QY 572 TTCAGTTTCCAATGATGAACATCATCAACGGTGGCGCTCAGCTGACTCCCGGTGTTGACG 631
 DB 422 TGCCTGTCGCGATGATGAACATCATCTCAACGGGGTGGCGACGCGGACCAACGTCGACA 481
 QY 632 TTCAAGNATTCATGATCGCTCCATCGGTGAGAGACCTTCTCTGAGGCTCTCCGCAACG 691
 DB 482 TCAGGAGTTTCATGATCGCGCCCATCGGGGCGGAGAGCTTTCGGGAGGCGCTGCGCTGG 541
 QY 692 GCGCGGAGGTTTCAACGCGACTGAAGTCCGTCATCAAGGAAAGGGCTGTCACCGGAC 751
 DB 542 GCGCGGAGGTTTCAACGCGACTGAAGTCCGTCATCAAGGAAAGGGCTGTCACCGGAC 601
 QY 752 TTGCGCATGAGGGCGGCTTCGCTCTCTTCCGCTCCACCGGTGAGGCTCTTTCACCTTA 811
 DB 602 TCAGTACGAGGCGGCTTCGCGCCCAACCTGGACAGCAACCGCGCGCTCGACCTGA 661
 QY 812 TCAGTAAAGCAATCGAGAGGCTGGCTTACCCCGAGGAGGACATCGCTCTTCTGCTGG 871
 DB 662 TCAGCGAAGCCATCACGAGGCGGCTTCACGCTGGGCGGAGGACATCGCTGCTGCTGG 721
 QY 872 ACCTTGCTTCTCTGAGTTCTTCAAGGACGCGCACTTACCACTTCGAAAGTGGCGGCACT 931
 DB 722 ACCTGCGCGCACCGAGTTCTACGCGACGCGGCTTACAGTTTCGAGGGTTCGAGCGCA 781
 QY 932 CCGCAGCTGAGATGGCAACGTTTACGCTGAGCTCGTTGACGCTACCCATCTCTCTCA 991
 DB 782 GCGCGGAGGAGTGGCGGCTTACTACACCGAGCTGGTTCGAGTCTTACCCGCTGCTGCA 841
 QY 992 TCAGGACCCACTCGAGGAGATGACTGGGAGGTTTACCAACCTCACCGCAACCACTCG 1051
 DB 842 TCAGGACCCGCTCAGCGAGGAGACTGGGCGGCTGGAAGGCGCTCACCGAGTCCCTCG 901
 QY 1052 GCGACAAGGTTTCAAGTTCGTTGGCGGAGCTTCTTCGTCACCAACCTGAGCGCTGAAGG 1111
 DB 902 GCGACCGGTCGAACCTGCTCGGCGAGCACTGTTTCGTCACCAACCTGAGCGCTGAGC 961
 QY 1112 AGGCGATCGCTAAGAGGCTGCGCAACTCCATCTCTGTTAAGGTGAACAGATCGGTACCC 1171
 DB 962 GCGGTATCGACGAGGCGCTGCGCAACTCGCTGCTGTTCAAGGTGAACAGATCGGTACGC 1021
 QY 1172 TCACCGAGACTTCGACGCTGTCGATGCTCACCGGCGAGGCTACACCTTCATGATGT 1231
 DB 1022 TCAGCGAGACTTTTGGACGCGGTGTCGCTGGCGGAGCGCAACCGGATGATCA 1081
 QY 1232 CCACCGTTCCGCTGAGACCGGAGGACACCACTTGTGCTGACCTCGGAGTTGCACTCACT 1291
 DB 1082 GCGACCGGTTGCTGAGAGACCGGAGACCACTGCTGCGGAGCTCGCGGAGTGGCGGAG 1141
 QY 1292 GTGCGCAGATCAAGACTGCTGCTCCAGCAGCTTCCGACCGTGTGCGAAAGTACCAACGAG 1351
 DB 1142 CCGGCGAGATCAAGACGCGGTGCGCGCGCGAGCGGCTGCGCAAGTACCAACGAG 1201
 QY 1352 TTCTCCGATTCGAGAGCTGCTTGGGCGACGCGCGGCTGTACGAGGTTGCGAGCGCATTC 1411
 DB 1202 TGCTGGGATTCGAAGAAGAGCTCGACGACGCGCGCTCTACGCGGAGCAACGCGTTTC 1261
 QY 1412 CAGCTTT 1418
 DB 1262 CCGCTTT 1268

RESULT 12
 ACA38079
 ID ACA38079 standard; DNA; 1287 BP.
 XX
 AC ACA38079;
 XX
 DT 19-JUN-2003 (first entry)

XX DE Prokaryotic essential gene #19736.

XX KW Antisense; ds; prokaryotic essential gene; cell proliferation;

XX KW drug design; gene.

XX OS Mycobacterium avium.

XX PN WO20027183-A2.

XX PD 03-OCT-2002.

XX PF 21-MAR-2002; 2002WO-US009107.

XX PR 21-MAR-2001; 2001US-00815242.

XX PR 06-SEP-2001; 2001US-00948993.

XX PR 25-OCT-2001; 2001US-0342923P.

XX PR 08-FEB-2002; 2002US-00072851.

XX PR 06-MAR-2002; 2002US-0362699P.

XX (ELIT-) ELITRA PHARM INC.

XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;

XX Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX WPI; 2003-029926/02.

XX P-PSDB; ABU34209.

XX New antisense nucleic acids, useful for identifying proteins or screening

XX PT for homologous nucleic acids required for cellular proliferation to

XX PT isolate candidate molecules for rational drug discovery programs.

XX Claim 14; SEQ ID NO 25949; 1766pp; English.

XX The invention relates to an isolated nucleic acid comprising any one of

XX the 6213 antisense sequences given in the specification where expression

XX of the nucleic acid inhibits proliferation of a cell. Also included are:

XX (1) a vector comprising a promoter operably linked to the nucleic acid

XX encoding a polypeptide whose expression is inhibited by the antisense

XX nucleic acid; (2) a host cell containing the vector; (3) an isolated

XX polypeptide or its fragment whose expression is inhibited by the

XX antisense nucleic acid; (4) an antibody capable of specifically binding

XX the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular

XX proliferation or the activity of a gene in an operon required for

XX proliferation; (7) identifying a compound that influences the activity of

XX the gene product or that has an activity against a biological pathway

XX required for proliferation, or that inhibits cellular proliferation; (8)

XX identifying a gene required for cellular proliferation or the biological

XX pathway in which a proliferation-required gene or its gene product lies

XX or a gene on which the test compound that inhibits proliferation of an

XX organism acts; (9) manufacturing an antibiotic; (10) profiling a

XX compound's activity; (11) a culture comprising strains in which the gene

XX product is overexpressed or underexpressed; (12) determining the extent

XX to which each of the strains is present in a culture or collection of

XX strains; or (13) identifying the target of a compound that inhibits the

XX proliferation of an organism. The antisense nucleic acids are useful for

XX identifying proteins or screening for homologous nucleic acids required

XX for cellular proliferation to isolate candidate molecules for rational

XX drug discovery programs, or for screening homologous nucleic acids

XX required for proliferation in cells other than *S. aureus*, *S. typhimurium*,

XX *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target

XX prokaryotic essential genes. Note: The sequence data for this patent did

XX not form part of the printed specification, but was obtained in

XX electronic format directly from WIPO at

XX ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 1287 BP; 214 A; 424 C; 475 G; 174 T; 0 U; 0 Other;

Query Match 41.08; Score 647; DB 8; Length 1287;

Best Local Similarity 69.98; Pred. No. 1.2e-153;

Matches 888; Conservative 0; Mismatches 380; Indels 3; Gaps 1;

151 GTGGCTGAAATCATGCACGATTTCGCTCGCGAAATTCGACTCCCGCGGTAAACCAACC 210

1 GTCCGATATCGACGAGTCCGGGCCCGCGAGATCCTCGACTCCCGCGGTAAACCAACA 60

211 GTCGAGCGACGAGGTTTTCTTGGATGACGGTTCCACCGTGTCCAGGTGTTCATCCGGC 270

61 GTCGAGGTGAGATCGCCCTGACCGACGGAACGTTGCCCCGCGCGGGTCCGCTCCGGG 120

271 GCATCCACCGCGCTCCACGAGGCTCATGAGCTGCGTGACGGTGGCGATCGTACTCTGGC 330

121 GCCTGACCGGTGAGCAGCAGGCTGCGAATCGCGGACCGCGGGGAGCGGTACGGCGGC 180

331 AAGCGGTTTTTGAAGGCAGTTGAAAACGTCACGAAGAAATCGCGCAGCAGCTCGCTGGC 390

181 AAGGGGTGCAGAAAGGGGTGAGGGCGTCTGAGCAGATCGCCCGCGGTGATCGGG 240

391 CTAGAGGTGACGATCAGCGCCCTCATCGACGAAGCAATGATCAAGCTTGTATGGCACCGCC 450

241 CTGAACGCGCAGCAGCAGCGCTGGTGGACAGCGCTCGTCTGACCTGGACGGCACCCCG 300

451 ACAAAGTCCCGCTGGGTGCAAAACGCAATCCTTGGTGTTCATGGCTGTTCGAAAGGCT 510

301 GACAAGTCCGAGGCTGGCGGCAACGCGATCCTGGCGGTGTCTGGCGCGTGGCCAAAGCG 360

511 GCTGCTGATTCGCGAGCGCTCCCACTGTTCCGCTACATCGGTGGAACCAACGACACAGTT 570

361 GCCGCGATTCGCGGAGTTGCCGCTGTTCCGCTACTCTGGCGGGCCCAACGGGCACATC 420

571 CTTCCAGTTCCAATGATGAACATCATCAACGCTGGCGCTCACGCTGACTCGGTGTTGAC 630

421 CTGCCGTGCGGATGATGAACATCCTCAACGGCGGTGCGCACGCCGACACCGCGCTGCAC 480

631 GTTCAGGATTCATGATCGCTCCAATCGGTGACAGACCTTCTCTGAGGCTCTCCGCAAC 690

481 ATCCAGAGTTTCATGTTGTCGCGCGATCGCGCGCGCGAGTTTCGCGGAGGATTCGCGTGG 540

691 GCGCGCGAGGTCTTACCACGACCTGAAGTCCGTCATCAAGGAAAGAGGCTCTGCACCGGA 750

541 GGTGCGGAGGTGTATCACTCGCTGAAGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 600

751 CTTGGCGATGAGGGCGGCTTCGCTCTTCGCTCGGCTCACCGCTGAGGCTCTTGACCTT 810

601 CTGGCGGACGAGGGCGGCTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTG 660

811 ATCGTTAAGGCAATCGAGAGGCTGGCTTCAACCCAGGCAAGGACATCGCTCTTGCTCTG 870

661 ATCGCGCGGCGCATCGAATCGCGCGGCTTCAAACTCGGACCGACGTCGTCGTCGTCGTCG 720

871 GAGCTTGTCTTCTCTGAGTTCTTCAAGGACGGCAC---CTACCACCTTCGAAAGTGGCCAG 927

721 GACGCGGCGCCACCGGAGTTCTACAGCGACGGCACCGGCTACAAGTTTCGAGGGCAGCAC 780

928 CATTCGCGAGCTGAGATGGCAAAAGTTTACGCTGAGTCTGTTGACGGGTACCCCAATCGTC 987

781 CGCACCGCGGAGCAGATGGCCGAGTTCTACGCGCGGCTGCTCGCGCGGTATTCGCTTGGTG 840

988 TCCATCGAGGACCCACTGACGAGGAGATGACTGGGAGGTTTACACCACTTCACCGCAACC 1047

841 TCCATCGAAGACCCGCTGTCGAGAGATGATGGACCGCTGGGGCGGCGCTGACCGCGTCG 900

1048 ATCGCGGACAAAGTTTCAGATCGTTGGCGACGACTTCTTCGTCACCAACCCCTGAGCGCTG 1107

901 ATCGCGGACCCGGTGCAGCTCGTCGCGGACGACGCTCTTCGTCGCAACCCCGAACGCGCTG 960

1108 AAGAGGGGATCGTGAAGAGGCTGCCAATCCTATCTGTTTAAAGTTGAACCATCGGT 1167

961 GAAGAGGGTATCGAAGAGGGCGTCGCAAAATGCGTTGCTGAAGGTGAATTCAGATCGGC 1020

1168 ACCTTACCGAGACCTTTCGACGCTGTCGACATGGCTCACCGCGCAGGCTACACCTCCATG 1227

1021 ACCTGACCGAGAGCTGGACGCGCTCGCTGGCCACACACAGCGGCTACCGCACGATG 1080

1228 ATGTCCACCGCTTCGCGTGAGACCGAGGACACCAACCATTCCTGACCTTCGAGTTGCACTC 1287

Db 1081 ATGAGCCACCGCAGCGGTGAACCGAGAGACACACAGATCGCCGACCTGGCGGTGGCCGTC 1140
 Qy 1288 AACTGTGGCCAGATCAAGACTGTGTCCAGCAGCGTTCCGACCGTGTCCGAAAGTCAAC 1347
 Db 1141 GGCAGCGGCAGATCAAGACCGGAGCGCGGCCGCGAGCGGTGGCAAGTACAC 1200
 Qy 1348 CAGCTTCTCGCATCGAGCAGCTGCTGGCGACCGCGCGGTCTACGAGGTTCGAGCGCA 1407
 Db 1201 CAGCTGTGCGGATCGAGGAGCGCTCGCGACGCGCGCGCTTACGCGCGGCACTGGCC 1260
 Qy 1408 TTCCACGCTT 1418
 Db 1261 TTCCCGGTT 1271

RESULT 13
 AAS59564/c
 ID AAS59564 standard; DNA; 30159 BP.
 XX AAS59564;
 XX DT
 XX 13-FEB-2002 (first entry)
 XX Propionibacterium acnes immunogenic protein encoding DNA #59.
 XX SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
 KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 KW dermatological; osteopathic; neuroprotectant; ds.
 XX Propionibacterium acnes.
 OS WO200181581-A2.
 PN 01-NOV-2001.
 PD 20-APR-2001; 2001WO-US012865.
 PF 21-APR-2000; 2000US-0199047P.
 PR 02-JUN-2000; 2000US-0208841P.
 PR 07-JUL-2000; 2000US-0216747P.
 XX (CORI-) CORIXA CORP.
 PA Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
 PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
 XX WPI; 2001-616774/71.
 XX Propionibacterium acnes polypeptides and nucleic acids useful for
 PT vaccinating against and diagnosing infections, especially useful for
 PT treating acne vulgaris.
 XX Claim 1; SEQ ID NO 59; 1069pp; English.

XX Sequences AAS59506-AAS59804 represent DNA molecules encoding
 CC Propionibacterium acnes immunogenic polypeptides. The proteins and their
 CC associated DNA sequences are used in the treatment, prevention and
 CC diagnosis of medical conditions caused by P. acnes. The disorders include
 CC SAPHO syndrome (synovitis, acne, pustulosis, hypertosis and
 CC osteomyelitis), uveitis and endophthalmitis. P. acnes is also involved in
 CC infections of bone, joints and the central nervous system, however it is
 CC particularly involved in the inflammatory lesions associated with acne
 CC vulgaris. A method for detecting the presence or absence of P. acnes in a
 CC patient comprises contacting a sample with a binding agent that binds to
 CC the proteins of the invention and determining the amount of bound protein
 CC in the sample. The polypeptides may be used as antigens in the production
 CC of antibodies specific for P. acnes proteins. These antibodies can be
 CC used to downregulate expression and activity of P. acnes polypeptides and
 CC therefore treat P. acnes infections. The antibodies may also be used as
 CC diagnostic agents for determining P. acnes presence, for example, by
 CC enzyme linked immunosorbent assay (ELISA). This sequence encodes the
 CC polypeptides shown in AAU54183-AAU54498 and AAU67555-AAU67557. Note: The

CC sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 30159 BP; 6067 A; 9304 C; 8907 G; 5878 T; 0 U; 3 Other;

Query Match 40.6%; Score 640.8; DB 4; Length 30159;
 Best Local Similarity 69.2%; Pred. No. 1.3e-151;
 Matches 889; Conservative 0; Mismatches 392; Indels 3; Gaps 1;

Qy 152 TGGCTGAATCATGCACGTATTCTGCTCGGAAATTCGCACTCCGCGGTAAACCAACCG 211
 Db 1539 TGGCAACCATCGAATTTCGAAAGCCGTCGAGATTCCTCGATTCCCGCGCAACCGACCG 1480
 Qy 212 TCGAGGCAGAGGTTTTCTCGATGACGTTTCCACGCTGTCGACAGGTGTTCCATCCGCGCG 271
 Db 1479 TTGAGGTCGAGATGATCTCGACGACGCGACCCAGGCCGCGCTGCGGTTCTTCGGGCG 1420
 Qy 272 CATCCACCGCGTCCACGAGGCTCATGAGCTGCGTGACGG---TGGCGATCGCTACCTGG 328
 Db 1419 CTTTCGACCGGTCAAGTTCGAGGCGGTTGAGCTGCGTGACGCGGATAAGAAGCGTTACTCCG 1360
 Qy 329 GCAAGGCGGTTTTGAAGGCGAGTTGAAACCGTCAACGAAAGAAATCGGACGAGCTCGCTG 388
 Db 1359 GTAAGGGTGTCTCAAGGCTGTTGAGAACTGTTCAACGAGAGATCGCCGAGAGGTGCTCG 1300
 Qy 389 GCCTAGAGGCTGACGATCAGCGCTCATCGACGAAAGCAATGATCAAGCTTCATGCGACCG 448
 Db 1299 GCTGTGACGCGAGTGACGCGCATTTATCGACCATGATCGAGCTCGACGATCCG 1240
 Qy 449 CCAACAGTCCCGCTGGGTGCAACGCAATCTTTGGTGTGTTTCCATGGCTGTTGCAAGG 508
 Db 1239 ACAACAAGGCAAGTTGGGTGCTAAACGCCATCTTTGGTGTCTCCCTGGCTGCTCATG 1180
 Qy 509 CTGCTGCTGATTCGCGAGGCTCCCACTGTTCGCTACATCGGTGAGCAACGACACG 568
 Db 1179 CTGCTGCTGACTGTGCGGAGCTCCCGCTGTACCAAGTACCTCGCGGAGACCGAACTCCACG 1120
 Qy 569 TTCTTCCAGTTTCCAAATGATGAACATCATCAACGTCGCGCTCACGCTGACTCCCGTGTG 628
 Db 1119 TGCTGCCGTTTCCAATGATGAACATCTCAACGTTGGTGGCCCGCGATTTCCGACGTTG 1060
 Qy 629 ACCTTCAGGAATTCATGATCGCTCCAATCGGTGACAGACCTTCTCTGAGGCTCTCCGCA 688
 Db 1059 ACATCAGGAGTTTCATGATGTCGCCGATCGCGCGGAATCTTTAAGCAGGCTACGAGT 1000
 Qy 689 ACGGCGGAGGCTTACCACGACTGAAGTCCGTCAATCAAGGAAAGGGCTGTGTCACCG 748
 Db 999 GGGGCGCTGCGGTTTACCACCTCCCTCAAGAAAGGTTCTCAAGGACAGGGCTTGGCTACCG 940
 Qy 749 GACTTGGCGATGAGGGGCTTCGCTCTTCGTCGCTCCACCGTGAGGCTCTTGACC 808
 Db 939 GTCTGGGTGACGAGGGCGGTTTCGCCCCCAACCTTCCACAGCAACGCCGCGCTTTGACC 880
 Qy 809 TTATCGTTAAGGCAATCGAGAAGGCTGGCTTCAACCCAGGCAAGGACATCGCTTTGCTC 868
 Db 879 TCATCTTTAGCGCATCAAGCGCGCGCTTCGAGCGGGTGAAGACGTCGCCCTCGCCC 820
 Qy 869 TGAAGCTTGGCTTCTCTGAGTTCTTCAAGGACGCGACCTACCACTTGAAGGTGGCGCAGC 928
 Db 819 TTGATGTGCTGCTTCCGAATTTCTCGAGACGCGCAAGTACACGTTTCGAGGGCCAGGCTA 760
 Qy 929 ACTCCGCGCTGAGATGGCAACGTTTACGCTGAGCTGTTGAGCGGTACCAATCTGCT 988
 Db 759 AGACCTCGGCTGAGATGATCGCGTATCTACGAAGGCTTCATCGCAAGTATCCGCTGCT 700
 Qy 989 CCATCGAGGACCCACTCGAGGAAGATGACTGGGAGGTTACACCAACCTCACCGCAACCA 1048
 Db 699 CTATTGAGGATCCGTTGGACGAGGAGATTGGACGCGGTGGGCGGAGTTTCAAGAGAGC 640
 Qy 1049 TCGGCGACAAGGTTTCAAGATCGTTGGCGACGACTTCTTTCGTCAACACCCCTGAGCGCTGA 1108
 Db 639 TTGGTGAGAAGATCCAGATCGTTGGCGACGACTTGTTCGTCACTTAACCCGAGCGCTTG 580

QY 1109 AGGAGGCGATCGCTAAGAGGCTGCCAATCCATCCCTGGTTAAGTGAACAGATCGTA 1168
 Db 579 CCAAGGGAATCGAGACCAAGGCTGCCAAGCCCTCGCTCGTCAAGGTGAACAGATCGCT 520
 QY 1169 CCCTCACCGAGACCTTCGAGCGTGTGACATGGCTCACCGCGAGGCTACACCTCCCATGA 1228
 Db 519 CCCTTCGGAGACTATCGAGCGCTGAGCTGGCCATCGAAGCGGTACCGTTCGATGA 460
 QY 1229 TGTCCCAACCGTTCGGTGAAGCGAGGACCAACCAATTCGTGACCTCGCAGTTGCACTCA 1288
 Db 459 TGTGCGACCGCTTCGTGTGAGACCGAGGACCAACCAATTCGTGACCTCGCAGTTGCACTCA 1288
 QY 1289 ACTGTGCGGAGATCGAGCTGTGCTCCAGACGTTCCGACCGTGTGCGAAGTACAAAC 1348
 Db 399 CGACTGTGTCAGATTAAATCCGGTGTCCCGCGCGCGCGAGCGATCGCCAAAGTACAAAC 340
 QY 1349 AGCTTCTCCGATCGAGCAGCTGTTCGGCGACGCGCGCTACGACAGGTTCGAGCGCAT 1408
 Db 339 AGTTGCTGCGTATTCGAGGAGAACTGGCGGACTCGGCGGAGTACGCGGTGCTTCGCTT 280
 QY 1409 TCCAGCGCTTCAGGGCTAAATAA 1432
 Db 279 TCCGCGCGCTTCAGGCGCTGACTGA 256

RESULT 14

ACF64493/C
 ID ACF64493 standard; DNA; 30159 BP.
 XX AC ACF64493;
 XX 17-OCT-2003 (first entry)
 DE Propionibacterium acnes DNA contig sequence #59.
 XX
 XX Acne vulgaris; antiseborrheic; dermatological; antibacterial;
 KW immunostimulant; immune response; vaccine; ds.
 XX
 OS Propionibacterium acnes.
 XX
 PN W02003033515-A1.
 XX
 XX 24-APR-2003.
 XX
 XX 11-OCT-2002; 2002WO-US032727.
 XX
 XX 15-OCT-2001; 2001US-00978825.
 XX
 XX (CORI-) CORIXA CORP.
 XX
 XX Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;
 PI Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;
 PI Barth B, Vallie-Douglas J;
 XX
 XX WPI; 2003-381789/36.
 XX
 XX New Propionibacterium acnes polypeptides and polynucleotides encoding the
 PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
 PT or for stimulating an immune response specific for a P. acnes protein.
 XX
 XX Claim 1; SEQ ID NO 59; 1481bp; English.
 XX
 XX The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
 CC encoding a Propionibacterium acnes protein. The invention also relates to
 CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to
 CC immunogenic fragments of P. acnes polypeptides. The invention
 CC additionally encompasses expression vectors and host cells comprising a
 CC polynucleotide of the invention; antibodies against polypeptides of the
 CC polynucleotide; fusion proteins comprising a polypeptide of the invention; a
 CC method for stimulating an immune response specific for a P. acnes
 CC polypeptide and an isolated T cell population comprising P. acnes polypeptides,
 CC via this method; and a vaccine composition (comprising P. acnes polypeptides,

CC polynucleotides, antibodies, fusion proteins, T cell populations, or
 CC antigen-presenting cells that express the polypeptide); a method and kit
 CC for detecting or determining the presence or absence of P. acnes in a
 CC patient; and a method for inhibiting the development of P. acnes in a
 CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion
 CC proteins, T cell populations or antigen-presenting cells that express the
 CC polypeptides are useful for diagnosing, preventing or treating acne
 CC vulgaris, or for stimulating an immune response specific for a P. acnes
 CC protein. The polynucleotides can also be used as probes or primers for
 CC stulmatic acid hybridisation. The vaccine composition is useful for the
 CC stimulation of an immune response against P. acnes, or for treating acne,
 CC and the kit is useful for performing a diagnostic assay. The present
 CC sequence represents a P. acnes DNA contig which is specifically claimed
 CC in the invention. Note: The sequence data for this patent did not form
 CC part of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 30159 BP; 6067 A; 9304 C; 8907 G; 5878 T; 0 U; 3 Other;
 Query Match 40.6%; Score 640.8; DB 8; Length 30159;
 Best Local Similarity 69.2%; Pred. No. 1.3e-151;
 Matches 889; Conservative 0; Mismatches 392; Indels 3; Gaps 1;
 QY 152 TGGCTGAAATCATGCACGTATTGCTCGCGAAATTTCTCGACTCCCGGGTAACCCACCG 211
 Db 1539 TGGCAACCATCGAATTCATCGAAGCCGTGAGATCTCGATTCCCGCGCAACCCGACCG 1480
 QY 212 TCGAGGACAGAGTTTTCCTGGATGACGGTTTCCACGGTGTGCGAGTGTTCATTCGCGCG 271
 Db 1479 TTGAGGTGAGATGATCTCGACGACGGCACCCAGGCCGCGTTCCTTCCGCGCG 1420
 QY 272 CATCCACCGCGTCCACGAGGCTCATGAGTGGTGCAGCG--TGGCGATCGTACTCG 328
 Db 1419 CTTGACCCGCTCAGTTTCGAGGCGGTTGAGTGGTGCAGCGGATAAGAGCGTTACTCCG 1360
 QY 329 GCAAGGCGGTTTGAAGGCGAGTTTGAAGAAAGCTCAACGAAGAAATCGCGACGAGTCTCG 388
 Db 1359 GTAAGGGTGTCTCAAGGCTGTTGAGAACGTCAACGAGAAGATCGCGGAGGTTGCTCG 1300
 QY 389 GCCTAGAGGCTGACATGACGGCTCATCGACGAAGCAATGATCAAGTTGATGGACCG 448
 Db 1299 GCTGTGACGCGAGTGAGACGCGCATTTATCGACAGATCATGATCGAGTCTGACGGATCCG 1240
 QY 449 CCAACAAGTCCCGCTGGGTGCAACGCAATCTCTGTTTCCATGGCTGTGTCGAAGG 508
 Db 1239 ACAACAAGGCAAGTTGGGTGCTTAACGCCATCTTGGTGTCTCCCTGGCTGCTCATG 1180
 QY 509 CTGCTGCTGATTCGCGAGGCTCCCACTGTTCCGCTACATCGGTGACCAACGACACG 568
 Db 1179 CTGCTGCTGATTCGCGAGGCTCCCGCTGTACCACTACCTCGCGGACCGAATCCACG 1120
 QY 569 TTCTTCAGTTTCAATGATGAACATCATCAACGGTGGCGTCAAGTGTGATCGCGTGTG 628
 Db 1119 TGTGTCGCGTTCOAATGATGAACATCTCAACGGTGTGTCACGCGCGATTCGAGCGTTG 1060
 QY 629 ACCTTCAGGAATTCATGATCGCTCCCAATCGGTGACAGACCTTCTCTGAGGCTCTCCGCA 688
 Db 1059 ACATCCAGAGTTTCATGATTCGCGGATTCGCGCGGAATCTTTAAGCAGGCTTACAGT 1000
 QY 689 ACGGCGGAGGTTTACCAACGCACTGAAGTTCGATCAAGGAAAGGCGCTGTTCACCG 748
 Db 999 GGGGCGCTGCGGTTTACCACTCCCTCAAGAAAGGTTCTCAAGGACAAAGGGCTTGGTACCG 940
 QY 749 GACTTCGATGAGGCGGCTTCGCTCCCTTCGTCGGCTCCACCGTGGAGCTCTTGACC 808
 Db 939 GTCTGGTACGAGGCGGCTTCGCGCGGCAACCTTCCAGACACGCGCGCTTGGAC 880
 QY 809 TTATCTTAAAGCAATCGAAGGCTTGGCTTCAACCCAGGCAAGGACATCGCTCTTGCTC 868
 Db 879 TCATCTTACGCGCATCAAGGCGCGGCTTCGAGCGGGTAAGGACGTCGCCCTCGCC 820
 QY 869 TGAAGCTTCTGCTGAGTTTCAAGGACGGACCTTACCACTTCGAAAGTTGGCGACG 928

Db 819 TTGATGTCGCTGCTTCGGAATTTTCGAGGACGCAAGTACACGTTTCGAGGGCCAGGCTA 760
 Qy 929 ACTCCGAGCTGAGATGCGCAACGTTTACGCTGAGCTCGTTGACGCGTACCAATCGTCT 988
 Db 759 AGACCTCGGCTGAGATGATCGGCTACTACGAAGGCTTATCGCAAGTATCCGCTGCT 700
 Qy 989 CCATCGAGGACCCACTCGAGAAGATGACTGGGAGGTTTACAACTCACCAGCAACCA 1048
 Db 699 CTATTGAGGATCCGTTGCGACGAGGAGATTGGGACGCTTGGCCGAGTTTACGAAGAAGC 640
 Qy 1049 TCGGCGACAGGTTTCAGATCGTTGGCGACGACTTCTTCGTCACCAACGCTGAGCGCTGA 1108
 Db 639 TTGTTGAGAAGATCCAGATCGTTGGCGACGACTTGTTCGTCACTAACCCGAAGCGCTTG 580
 Qy 1109 AGGAGGCGATCGCTTAAGAAGGCTGCCAACTCCATCTCTGGTTAAGTGGAACAGATCGGTA 1168
 Db 579 CCAAGGGAATCGAGACCAAGGCTGCCAAGCGCTGCTCGTCAAGTGAACAGATTGGCT 520
 Qy 1169 CCTCACCGAGACTTCGACGCTGTCGACATGCTCACCGGCGAGGCTACACCTCCATGA 1228
 Db 519 CCTCTCGGAGACTATCGACGCGCTCGAGCTGGGCCCATCGCAACGGCTACCGCTGCATGA 460
 Qy 1229 TGTCCGACCGTTCGCTGACCGGAGGACACCACTTGTGCTGACCTCGAGTTGCATCA 1288
 Db 459 TGTGCGACCGTTCGCTGAGACCGAGGACCACTATCGCGGATCTCGCCGCTCGCGCTAT 400
 Qy 1289 ACTGTGGCCAGATCAAGACTTGGTCTCCAGCACGTTCCGACCGTTCGCAAGTACAAAC 1348
 Db 399 CGACTGCTGAGATTAAATCCGTTGCCCGCCGCGCGGAGCGATCCCAAGTACAAAC 340
 Qy 1349 AGCTTTCGCGATCGAGCAGCTGCTTGGCGAAGCGCGCGCTTACGCGCAGCGCAT 1408
 Db 339 AGTTGCTGCTATTGAGGAGGAATGCGGCGACTCGGCCGAGTACGCCGCTTCCGCTT 280
 Qy 1409 TCCGACGCTTTCAGGGCTAATAA 1432
 Db 279 TCCGCGCTTCCAGGCGCTGACTGA 256

RESULT 15

ABX09144

ID ABX09144 standard; DNA; 45190 BP.

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

This invention relates to a novel method for identifying essential genes for survival or virulence of mycobacteria species. The method comprises aligning the genomic sequence of a first mycobacterium species on a genomic sequence of a second mycobacterium species and selecting a polynucleotide sequence that is highly conserved in both genomes with no counterparts in other bacterial genomic sequences and that corresponds to an essential gene for the survival or virulence of mycobacterium species. The method of the invention is useful for detecting M. tuberculosis or M. leprae infection. The method reduces the number of potential new targets and protective antigens for new drugs and vaccine compositions to treat and prevent mycobacterial diseases, particularly tuberculosis and leprosy. The present sequence represents a Mycobacterial cosmid DNA sequence used in the method of the invention

Sequence 45190 BP; 7490 A; 14522 C; 15204 G; 7974 T; 0 U; 0 Other;

Query Match 40.1%; Score 632.6; DB 6; Length 45190;

Best Local Similarity 68.4%; Pred. No. 1.8e-149;

Matches 891; Conservative 0; Mismatches 409; Indels 3; Gaps 1;

140 AGGAGCCACAGTGGCTGAATCATGCACGTTATTCGCTCGGAAATTCGACTCCCGCG 199

19931 AGGAGAACCCAGTGCCTGATTCATCGACGAGTTAGGCCCGGAGAGATCCTCGATTCCCGCG 19990

200 GTAACCCAAACCCGCTGAGGAGAGGTTTCTCGATGACGCTTCCACGCTGTCGACAGTG 259

19991 GCACCCGAGCTGGAGTTCGAGTGGCGTTATCGACGGGACATTCGCCGGGCCCGCG 20050

260 TTCCATCCCGCGCATCACCGGCTCCAGGCTCATGAGCTCATGAGCTGCGTACGGTGGCGATC 319

20051 TGCCTGTCGGCGCTCGACCGGGGAGCACGAGCGCTGCGAGTTGCGCGACGCGCGCGATC 20110

320 GCTACTCTGGCAAGGCGTTTGAAGCAGTTGAAACGCTCAACGAGAAATCGCGACG 379

20111 GCTACGCGGCAAGGCGTGCAGGAGCGTGCAGGCTGTTCTTGTGAGATCGGCGCGG 20170

380 AGCTCGCTGGCTTAGAGGCTGAGCATCAGCGCTCATCGACGAGCAATCATCAAGTTG 439

20171 CCGTCATCGGACTCAACCGCGAGCAGCAGGATTTGTCGACAGGCGCTGGTGACCTAG 20230

440 ATGGCACCGCCAAAGTCCCGCTGGGTGTCACCAACGCAATCTTGTGTTTCCATGCTG 499

20231 ACGCCACCCCGACAAAGTCCCGCTGGCGGCAACGCGATCTTGGGTGTCTCGCTCGCTG 20290

500 TTGCAAGGCTGCTGCTGATTCGCGAGGCTCCACCTGTTCCGCTACATCGTGGACCA 559

20291 TTGCCAAGCGCGCGGAGATTTCGCGGAGCTGCGCTTGTTCCTGTTATGTCGGGGGCGCA 20350

560 ACGCACACGTTCTTCCAGTTCCCAATGATGAACATCATCAACGCTGGCGCTCACGCTGACT 619

20351 ACGCGCACATTCTGCGGTACCGATGATGAACATCTCAACGCGGCGGCGACACGCCGATA 20410

620 CCGGTGTGACGTTTCAGGAATTCATGATCGCTCCAATCGGTGCGAGACCTTCTCTGAGG 679

20411 CCGCTGTGACATTCAAGAGTTTATGTTGGCGCAATTTGGCGCGCCAGCTTCTGTCAGG 20470

680 CTCTCCGCAACGCGCGGAGTCTACCAACGCTGAGTCCGTCATCAAGGAAAGGCGCC 739

20471 CGTTGCGCTGGGCGCTGAGGTGTACCAACGCTCAAGTCCGCTTCTGCTGCTGAGGCGG 20530

740 TGTCCACCGGACTTGGCGATGAGGCGGCTTCCGCTCCTTCCGCTCGGCTCCACCGCTGAGG 799

20531 TGTCCACCGGCTTGGCGGAGGAGGCGGCTTCCGCGGATGTGGCGCGGACACCGCGG 20590

800 CTCTTGACCTTATCGTTAAGGCAATCGAAGAGCTGCTTACCCCGGCAAGGACATCG 859

20591 CGTTGACCTGATCAGCGCGGCGCATCGAGTCCGCGGCTTGGCGCGCGCGCGACGCTGG 20650

860 CTCTTGCTGAGAGCTTGTCTCTCTGAGTTCTTCAAGGACGCGAC---CTACACCTTCG 916

20651 CGCTGGCCCTTGGACGCGCGGCGCCACGAGTTCTTCCCGGACGCGCACCGGCTACGCTTTCG 20710

Identifying and selecting genes for survival or virulence of mycobacteria by a comparative genomic analysis of the sequences of Mycobacterium tuberculosis and M. leprae.

Disclosure; Fig 8; 874pp; English.

Qy	917	AAGTGGCCAGCACTCGGAGCTGAGATGCAACCGTTTACGCTGAGCTCGTTGACGCGT	976
Db	20711	AGGGCAACCCGTACCGCAGACGAGATGACCGAGTTCTACGGCGGCTGCTCGGCGCT	20770
Qy	977	ACCAATCGTCTCATCTGAGGACCCACTGCGAAGATGACTGGGAGGGTTACACCAACC	1036
Db	20771	ACCGCTGGTGTGATCGAAGACCCACTGTCCGAGACGATTGGACGGCTGGGCCCGCGC	20830
Qy	1037	TCACCGCAACCATCGGCGACAAGGTTTCAGATCGTTGGCGAGACTTCTTCGTCAACCAACC	1096
Db	20831	TGACGGCCTCGATCGGTGACCGGGTGCAATCGTCGGCGACGACATCTTTGTCAACCAATC	20890
Qy	1097	CTCAGCGCCTGAAGGAGGCGATCGCTAAGAGGCTGCCAATCCATCCCTGGTTAAGGTGA	1156
Db	20891	CCGAGCGGCTCGAGGAGGGGATCGAACGGGGCGTGGCAATGCGTTGCTGGTCAAGGTGA	20950
Qy	1157	ACCAGATCGGTACCCCTCACCGAGACCTTTCGACGCTGTTCGACATGCGTTCACCGCGCAGGCT	1216
Db	20951	ACCAGATCGGAGCGTTGACCGAGACACTCGACGGGTTCACGCTGGCTCACCCACGGCGGAT	21010
Qy	1217	ACACCTCCATGATGTCCACCGTTCCGGTGAGACCGGAGGACACCAATTCGTGACCTCG	1276
Db	21011	ACCGCAGATGATCAGTCACCGCAGTGGCGAGTGGCGAGACGGAGGACACCATGATCGCCGACCTCG	21070
Qy	1277	CAGTTGCACTCAACTGTGGCCAGATCAAGACTGCTGCTCCAGCACGTTCCGACCGGTGTCG	1336
Db	21071	CGGTGGCCATCGGACGGGCGAGATCAAGACGGGCGGCTGCTCGCAGTGAGCGGCTCG	21130
Qy	1337	CAAAGTACAACCAAGCTTCTCCGATCGAGCAGCTGCTTGGCGACGCGCGGCTTACGCGAG	1396
Db	21131	CAAAATACACCAAGCTGCTGGGATCGAAGAGGCGCTTGGGACGCGGCGCCGCTACGCGG	21190
Qy	1397	GTCGACGCGATTCCCAACGCTTTCAGGGCTAAATAAAGCGCT	1439
Db	21191	GCGACCTGGCATTTCCTCGGTTCGCGTGCAGACGAAATAGGT	21233

Search completed: September 28, 2005, 17:01:34
 Job time : 796.5 secs

GenCore version 5.1.6

	Query Match	25.7%;	Score 406;	DB 4;	Length 1522;
	Best Local Similarity	60.4%;	Pred. No. 2.1e-98;		
	Matches 786;	Conservative	0;	Mismatches 465;	Indels 51; Gaps 5;
QY	143	AGGCCACAGTGGCTGAAATCATGCAGTATTCTCGTCCGAAATTTCTCGACTCCCGCGGTA	202		
Db	30	AAGCAACAATGTGCGACATCAAGAGCGTCTTCGCGCGCGAGATCCTGGACTTCGTGGCA	89		
QY	203	ACCCAAACCGTTCGAGGCAGAGGTTTTCTGGATGACACGGTTCGCCAGGTGTTCGACAGGTGTC	262		

```
Db 90 ACCCCACCGTCGAGGTGACCTGACACACGAGAGGCGCTCTTCGG---CTGGCGGTCC 146
Qy 263 CATCCGGCGCATCCACGGCGTCCACGAGGCTCATGAGCTGCGTGAAGTGGCGA---TC 319
Db 147 CCTCTGGTGCCTCGACTGGCATCTATGAGGCTCGAGCTCCGTGACGGCGACAAGTCGG 206
Qy 320 GCTACCTGGGCAAGGGGTTTGAAGGAGTTGAAACGTTCAACGATCAAGAAATCGGCGAG 379
Db 207 GCTACCTCGGCAAGGGGTGCTGAAGGCGCTGAGAAAGTCAACAGATCCTGGCGCCCA 266
Qy 380 AGCTCGTGCCTAGAGGCTGACGATCAGGCGCTCATCGACGAGCAATGATCAAGCTTG 439
Db 267 AGTGATCGGCTCGAGCTCACAAAGCAGGCGAGATCGACAGGCTCATGCTCAGATCG 326
Qy 440 ATGGCACCGCAACAAGTCCCGCTGGGTGCAACGCAATCCCTTGGTGTTCATGGCTG 499
Db 327 ACGGCACTGAGAAACAAGACCCACTGGGCGCCAAAGCCATCTCTGGTGTGCTGATGTCGG 386
Qy 500 TTGCAAGGCTGCTGTGATTCGCGAGGCTCCCACTGTTCCGCTACATCGGT----- 552
Db 387 TCTGCGGTGGCGCGCGCTTTCAGGGGCGCTGCGCTCTTACCGCTACATCGCGAGCTGT 446
Qy 553 --GGACCAACGACACGTTTCTTCCAGTTCGAATGATGAACATCATCAACGCTGGCGCTC 610
Db 447 CGGGCAACAAGTCCGCCATCTGCTCTGCCCTGCTTCAAGCTCATCAACGGCGGTGAGC 506
Qy 611 ACCTGATCTCCGGTGTGAGTTCAGGAATTCATGATCGCTCCAAATCGGTGCGAGACCT 670
Db 507 ACGGGGCAACAAGCTCGCCATGCGAGGATTCATGATCTGCCCCACCGGGCGGACCTCGT 566
Qy 671 TCTCTGAGGCTCTCCGCAACGGCGCGAGGTCTACCAAGGCTGAGTCCGCTCATCAAGG 730
Db 567 TCCACGAGGCGCTCGCATGTCGCCCGCGAGAGCTACCAACAACCTGAAGCTCGTCAAGA 626
Qy 731 AAAAGGGC-----CTGTCCACCGGACTTGGCGATGAGGGCGGTTCGCTCCTTCCGTCG 784
Db 627 AGAGTAGGCATGAGCGCGACCAACGTGCGGCGAGGAGGTGGCTTCGCCCCCAACATCC 686
Qy 785 GCTCCACCGGTGAGGCTCTTGACCTTATGCTTAAGGCAATCGAGAAGCGTGGCTTACCC 844
Db 687 AGGCCAACCAAGGCGCTCGAGCTCATCGTTGAGGCGCATCAAGCAGGCGCGGTACACTG 746
Qy 845 CAGCGAGGACATCGCTCTTGCTCTGACGCTGCTTCTCTGAGTTCCTCAAGGACGGCA 904
Db 747 GCAAGATCGAGATCGGATGACGTTGCGCGCTCGTCTGTTGAGGACGCCAAGAGTCGA 806
Qy 905 CCTACCACTTCGAAGGTGGCCAGCACTCCGAGCTGAGA----- 943
Db 807 AGTAGCACTCGGCTTCAAGGTCCCGCGGACAGAAGACCCCGACATGCTGCTCTCGG 866
Qy 944 -----TGGCAACGTTTACGCTGAGCTGTTGAGCGGTACCGAATGCTCTCCATCG 994
Db 867 GCGAGGGCGCTGATCAAGCTCTACGAGGAGTGGACCTCGAAGTACCCGATCTGGTTCGATCG 926
Qy 995 AGGACCCACTCGAGGATGACGTTGGAGGGTTACACCAACTCACCGCAACCATCGCG 1054
Db 927 AGGACCCCTTCGACGAGGACGCTGGGCCACTTACACCAAGGTTCAACGAGCTGATCGCA 986
Qy 1055 ACAAGGTTTCAGATCGTTGGGACGACTTCTTCTGTCACCAACCTCGAGCGCTGAAGGAGG 1114
Db 987 ACAGGATCCAGATCGTGGCGGACGACCTGCTGGTTCACCAACCCCAAGCGCATCTCGAGG 1046
Qy 1115 GCATCGCTAAGAGGCTGCCAATCTCCATCTCTGGTTAAGGTGAACAGATCGGTACCTCA 1174
Db 1047 CCGCAACAAGAGGCGCTGCAACGCGCTGCTGCTCAAGCTCAACAGATCGGCTCGGTCA 1106
Qy 1175 CCGAGACCTTCGACGCTCGACATGGCTCACCGCGCAGGCTACACCTTCCATGATGTCCT 1234
Db 1107 GCGAGGCGCTCGAGGCGTGGCGCTGGGCCGCGAGGTCAACTGGGGCGTCAATGTTGTCG 1166
Qy 1235 ACCGTTCCGTTGAGACCGAGGACACCAATGCTGAGCTCGCAGTTGCACTCAACTGTG 1294
```

```
Db 1167 ACCGCTCGGGCGAGACTGAGGACGCGTTTCATCGCCGAACTCGTCTCGGCTCGGCTCG 1226
Qy 1295 GCCAGATCAAGACTGGTCTCCAGACGTTCCGACCGTTCGCAAGTACACACGAGTTC 1354
Db 1227 GCCAGATCAAGACTGGTGCCTCCCTCGCTCGAGGCGCTCGCAAGTACACACGAGTCC 1286
Qy 1355 TCCGATCGAGCAGCTGTTGGCGACGCGCGGCTCTACGCG 1396
Db 1287 TCCGATCGAGGAGGCTCGGCGCAAGGCCACTACGCG 1328

RESULT 2
BH770842/c
LOCUS
DEFINITION
  BH770842 2491 bp DNA linear GSS 01-MAY-2002
  LLMtag581 MG1363 Random Sequence Tag Library Lactococcus lactis
  subsp. cremoris genomic, genomic survey sequence.
ACCESSION
  BH770842
VERSION
  BH770842.1 GI:20373799
KEYWORDS
  GSS.
SOURCE
  Lactococcus lactis subsp. cremoris
  Lactococcus lactis subsp. cremoris
  Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
  Lactococcus.
ORGANISM
  1 (bases 1 to 2491)
  Bolotin,A., Ehrlich,S.D. and Sorokin,A.
  Studies of genomes of dairy bacteria Lactococcus lactis
  Sci. Aliments (2002) In press
  Contact: Sorokin A
  Genetique Microbienne
  INRA
  CRJ INRA, Domaine de Vilvert, 78352 Jouy en Josas cedex, France
  Tel: 33 1 34 65 25 16
  Fax: 33 1 34 65 25 21
  Email: sorokine@jouy.inra.fr
  Best homologue in strain IL1403 is enoA (100%)
  Class: shotgun
  High quality sequence start: 30
  High quality sequence stop: 2463.
  Location/Qualifiers
FEATURES
  source
    1..2491
    /organism="Lactococcus lactis subsp. cremoris"
    /mol_type="genomic DNA"
    /strain="MG1363"
    /sub_species="cremoris"
    /db_xref="taxon:1359"
    /clone_lib="MG1363 Random Sequence Tag Library"
    /notes="Vector: pSGMU2; Site 1: SmaI; Library of
    chromosomal fragments of L.lactis strain MG1363 was
    prepared by partial AluI digestion or by sonication."
ORIGIN
  Query Match 25.3%; Score 399.2; DB 8; Length 2491;
  Best Local Similarity 60.3%; Pred. No. 1.7e-96;
  Matches 722; Conservative 0; Mismatches 448; Indels 24; Gaps 3;
  Qy 259 GTTTCATCCGGCGCATCCACGGCGCTCCAGAGGCTCATGAGCTCGGTGCGGCGA- 317
  Db 2485 GTACCTTCAGTCTCTACTGCTGACACGAGCGTTGAACCTTCGTGATGCGCAAA 2426
  Qy 318 --TCGTTACCTGGGCAAGGCGTTTGAAGGAGTTGAAAACGTCAACGAGAAATCGGC 375
  Db 2425 TCTCGCTACACGAGCTTGGTACTCAAAAAGCTGTTGACAACTGAACCAACATCATCGCT 2366
  Qy 376 GAGCAGCTCGCTGGCTAGAGGCTGAGATCAGCGCTCATCGACGAAAGCAATGATCAAG 435
  Db 2365 GAAGCTATCATCGTTATGAGTTACTGACCAACAGCTATTGACCGTGAATGATCGCT 2306
  Qy 436 CTTTATGGGACCGCCAAAGTCCCGCTCGGTCGAAACCGCAATCTCTTGTGTTTCATG 495
  Db 2305 CTTGACGCTACTGAAAACAAAGGTAAATTCGGAGCTAACGCTATTCTTGTGTTTCATC 2246
  Qy 496 GCTGTTGCAAGGCTGCTGCTGATTCCGCGAGGCTCCCACTGTTCCGCTACATCGGTGA 555
```

D	b		2245	GCTGCTGCTCGTCTGTGATGAACATTGGTGTTCCACTTTTACAACACTACCTTGGCGGA	2186
Q	y		556	CCAAACGCACAGTTCCTTCAGTTCCTCAATGATGAACATCATCAACGGTGGCGTCA	615
D	b		2185	TTCACGCTAAAGTATTGGCCAACTCCTCAATGATGAACATCATCAATGGTGGTCTCACTCA	2126
Q	y		616	GACTCCGGTGTGACGTTTCAGGAATTCATGATCGCTCCAATCGGTGAGAGACCTTCTCT	675
D	b		2125	GAGCCCCATCGCTTTTCCAAGAATTCATGATCGTACCAGTTGGTGACCTCATATTCAA	2066
Q	y		676	GAGGCTCTCGCAACGCGCGGAGGTCTACCACGACTGAAGTCGGTCATCAAGGAAAAG	735
D	b		2065	GAAGCGCTTCGTTGGGGTGTGTAATCTTCCACGCTTTAAGAAAAATCTTAAGCTCGT	2006
Q	y		736	GGCTGTCCACCGGACTTGGCGATGAGGGCGGCTTCGCTCTTCCTCGGTGCCACCCCGT	795
D	b		2005	GGACTTGAACAGCTGTCGGTGACGAAGGTGGATTTCGCTCCTAAATTCGACGGAACTGAA	1946
Q	y		796	GAGGCTCTTACCTTATCGTTTAAGGCAATCGAGNAGGCTGGCTTCACCCCAGGCAAGGAC	855
D	b		1945	GACGGTGTAGAAACTATCTTTAAAGCAATCGAAGCAGCTGGTTACAAGCTGGTGAGAT	1886
Q	y		856	ATCGCT--CTTGCTCTGGAGCTTTCCTCTGTGAGTTCCTCAAGGACGCACTACCAC	912
D	b		1885	GGCGTTATGATCGGTTTCGACTGTGCATCATCAGAAATCTACGAAAACGGTGTTCAGAC	1826
Q	y		913	TTCT-----GAAGGTGGCGCAGCACCTCCGACGCTGAGATGGCAACGTT	954
D	b		1825	TACACTAAATTCGAAGGTGAAGCGGTAAATACTTTTCAGTTCGTAACAAGTTGACTAC	1766
Q	y		955	TAGCTGAGCTCGTTTAGCGGTACCCAAATCGTCTCCATCGAGGACCCACCTGCAGGAAGAT	1014
D	b		1765	CTTGAAGAACTCGTTTCTAAATACCCAAATCATCACTATTGAAGATGGTATGGACGAAAC	1706
Q	y		1015	GACTGGAGGTTACACAACTCACCGCAACCATCGCGCACAGGTTACAGATCGTTGGC	1074
D	b		1705	GACTGGGATGGATGGAAAAATCCCTTACTGAAACGCTTGGTAAAAAGTTTCAACTCGTTGGT	1646
Q	y		1075	GAGCACTTCTTCGTACCAACCCCTGAGCGCTGAAGGAGGCGCATCGCTAAGAAGGCTGCC	1134
D	b		1645	GACGACTTCTTCGTACAAACACTAAATACCTTGAACGTGGTATCCGTGAAATGCTTCA	1586
Q	y		1135	AATCCCATCTGTTTAAGGTGAACACAGATCGGTACCCCTCACCGAGACCTTCGACGCTGC	1194
D	b		1585	AAGCCTATCTTGATCAAAGTTAACCAATCGGTACTTTTGACAGAAACTTTTGAAGCTATT	1526
Q	y		1195	GACATGCTCACCGCGAGGCTACACCTCCATGATGTGCCACCGTTCGGGTGAGACCGAG	1254
D	b		1525	GAATGGCTAAAGAAGCTGGTTTTACAGCAATCGTATCTCACCGCTTCAGGTGAAACTGAA	1466
Q	y		1255	GACACCACTATGCTGACCTCGCAGTTGCACTCAACTGTGGCAGATCAAGACTGGTGCT	1314
D	b		1465	GATTCAACAATCTCAGACATCGCTGTTTGCAACTAACGCTGGTCAATCAANAATGGTTCA	1406
Q	y		1315	CCAGCAGTTCGACCGGTGTCGAAAAGTACAACAGCTTCTCCGCAATCGAGCAGCTGCTT	1374
D	b		1405	CTTTCAGTACAGACCGTATGGCTTAATACAACCAATTGCTTCGTATTGAAGACCAATTG	1346
Q	y		1375	GGCGACCGCGGCTTACGCGAGGTCGACGGCATTTCCACAGCTTTTCAGGCGCTAA	1428
D	b		1345	GCTGAAGTTGCTCAATAAAGGCTTTAAAGCATTTTACAACCTTTAAAAATAA	1292

RESULT 3

CR685174	CR685174	1626 bp	mRNA	linear	HTC 19-AUG-2004
DEFINITION	Tetraodon nigroviridis full-length CDNA.				
ACCESSION	CR685174				
VERSION	GI:51183081				
KEYWORDS	HTC; CDNA; full-length; Tetraodon nigroviridis.				
SOURCE	Tetraodon nigroviridis				
ORGANISM	Tetraodon nigroviridis				
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				

REFERENCE	Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes; Tetraodontoidea; Tetraodontidae; Tetraodon.
AUTHORS	1 (bases 1 to 1626)
TITLE	Genoscope.
JOURNAL	Direct Submission
COMMENT	Submitted (10-AUG-2004) Genoscope - Centre National de Sequencage : 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr) The sequences are based on single pass reads. More information available at http://www.genoscope.cns.fr/tetraodon.
FEATURES	Location/Qualifiers
source	1..1626
ORIGIN	/organism="Tetraodon nigroviridis" /mol_type="mRNA" /db_xref="taxon:99883" /tissue_type="Eyes"
Query Match	24.6%; Score 388.4; DB 3; Length 1626;
Best Local Similarity	61.0%; Pred. No. 1.3e-93;
Matches	799; Conservative 0; Mismatches 451; Indels 60; Gaps 8;
Qy	160 ATCATGCACGTATTTCGTCGCGAAATTTCTCGACTCCGCGGTAAACCAACGTCGTGAGGCA 219
Db	110 ATCTGAAGATCACGCTCGGGAATATTTCGACTCCGCGGCAACCCACCGCTGGAGTTC 169
Qy	220 GAGGTTTTCGTGATGACGGTTCCACGGTGTGCGAGGTGTTCCATCCGGCGCATCCACC 279
Db	170 GACCTGTACCAAGAAAGGTCTGTTTCAGAGCTGCG---GTCCCCAGCGGCGCTTCACC 226
Qy	280 GGCCTCCACGAGGCTCATGAGCTCGGTGACGGTGGCGA---TCGCTACCTGGGCAAGGGC 336
Db	227 GGCATCTACGAGGCCCTGAGCTCCGCGACACGACAAACCCGCTACATGGGCAAGGT 286
Qy	337 GTTTTGAAGCGATTGAAACGTTCAACGAGAAATTCGGCGACGAGCTCGTGGCTTAGAG 396
Db	287 GTCTCTAAAGCTGTTGAGATATCAATATAAATAATTGCACCTGCACTGGTTGGCAAGGAC 346
Qy	397 GCTGAGCATCAGGCC-----TCATCGACGAAGCAATGATCAAGCTTGATGGCACCGCC 450
Db	347 GTCAAGCTTCTGACACAGGCAAAATTCGAACGCTGATGCTGGAGATGGACGGCACCGAC 406
Qy	451 AACAACTCCCGCTGGGTGCAACGCAATCCTTTGGTGTTCATGGCTGTGTCAAAGGCT 510
Db	407 AACAACTTAAGTTGGGGCCACGCACTCTGGCGGTGTCCCTGGCTGTGTGCAAGGCT 466
Qy	511 GCTGCTGATTCGACAGGCTCCCACTGTTCCGCTATCATCGGTGGACCAACGCG----- 563
Db	467 GGTGCACGAGAGAAGGCGTGGCCACTCTACCGCCACATCGCCGACCTGGCGCGCAACCCC 526
Qy	564 --ACAGTTCCTTCAGTTTCCAATGATGAACATCATCAACGGTGGCGCTCAGCTGACTCC 621
Db	527 CAAGTCATCTCCCGCTTCCCGCTTTCAACGTCATCAACGGCGGCTCCCAACGCAAGAAC 586
Qy	622 GGTGTTTGACCTTCAGGAATTCATGATCGCTCCAAATCGGTGCAGAGACCTTCTCTGAGGCT 681
Db	587 AAGCTGGCCATCAGAGATTTCATGATCTCGCCGGTGGAGCCAGACAGCTTCAAGAGGCC 646
Qy	682 CTCGCAACGGCGCGGAGGTCTACCAACGCACTGAAAGTCCGTCATCAAGGAAAAGGGC--- 738
Db	647 ATCGCATCGGCGCCGAAGTCTACCAACACTTGAAGAACGTCATCAAGGAGAGATACGGC 706
Qy	739 ---CTGTCAACGGCACTTGGCGATGAGGGCGGCTTCGCTCTTCGTCGGCTCACCCGT 795
Db	707 AAGGACGCCCAACAGCTGGGAGACGAGGGAGGCTTTTCCGCCCAACATCTCTGGAGAACAG 766
Qy	796 GAGGCTCTTGACCTTATCGTTAAGGCAATCGAAGGCTGGCTTTCACCCAGGCAAGGAC 855
Db	767 GAAGCTCTGAGCTGCTGAAGAACGCCATCGTCAAGGCGGGCTATAC-----CGACNAG 820
Qy	856 ATGGCTCTTGCTCTGGACGTTGCTTCTCTGAGTTCTTCAAGGACGCGCACTTACCACCTC 915

```

Db      821  ATTGTGATCGGCATGGACGTGGTGGCTCTGAGTTCTACAAGGGCGGCAAGTACGACCTG 880
Qy      916  GAA-----GGTGGCCAGACATCCCGACGTGAGATGGCAACGTT 954
Db      881  GACTTCAAGTCTCCGACGACCCGCGCGTATCATCTCTCCGAGCAGTGGCGACCTC 940
Qy      955  TAGCTGAGTCTGTTGAGCGGTACCCAAATCGTCTCCATCGAGAACCCACTGCAAGGAAT 1014
Db      941  TACAGGGGCTTCGTCAAAGATTACCCAGTGGTGTCCATCGAGACCCCTTTTGACAGGAC 1000
Qy      1015  GACTGGGAGGTTTACACCACTTACCGCAACCATCGGCGACAAGGTCAGATCGTTGGC 1074
Db      1001  GACTGGGAGGCTGGACCAATTTACAGCCAGCACCCAGC-----ATTCAAGTGGTGGG 1054
Qy      1075  GAGGACTTCTTCGTACCAACCTCTGAGCGCTGAAGGAGGCGATCGTAAAGAGGCTGCC 1134
Db      1055  GACGACCTCAGGTACCGGACCCCAACGATCGCCAAAGGTTGGCCCAAGATCTCTGC 1114
Qy      1135  AACTCCATCTGTTAAGGTGAACAGATCGGTACCTCTCCAGAGCATTTCGACGTGTC 1194
Db      1115  AACTGCCGTGTCTCAAAGTCAACAGATCGGTCTGGTCAACGAGTCCCTGCGAGCATGC 1174
Qy      1195  GACATGCTCACCGCGAGGTACACCTCCATGATGTCGCCAGCTTCGGTGAAGCGAG 1254
Db      1175  AAGATGCCCCAGAGCAACCGCTGGGGCGTGATGTCAGCCATCGCTCCGGGGAGACGGAG 1234
Qy      1255  GACACCACTATGTCGACCTCGCAGTTGCACTCAACTGTGTGGCCAGATCAAGACTGTGCT 1314
Db      1235  GACACCTTCTGCTGACCTGGTGTGGTCTGTGACACCGGACAGATCAAGACGGGCGCT 1294
Qy      1315  CCAGCAGTTCGACCGACGTGTGCAAGTACAAACAGCTTCTCCGATCGAGCAGTGTCTT 1374
Db      1295  CCGTGGCGCTCCGAGCGCTGGCCAAAGTACAAACAGCTGTCTCGGATCGAGGAAGCTG 1354
Qy      1375  GGCAGCGCGCGCTACGAGGTCGCGAGCGATCCAGCTTCCAGCTTTCAAGG 1424
Db      1355  GGCAGCAGGCGCGCTTCGCGGCAAGAACTTCAGGCACCCCATCTGAGCG 1404

RESULT 4
CR726296
LOCUS      1643 bp      mRNA      linear      HTC 19-AUG-2004
DEFINITION Tetraodon nigroviridis full-length cDNA.
ACCESSION CR726296
VERSION   GR51244547
KEYWORDS  HTC; cDNA; full-length; Tetraodon nigroviridis.
SOURCE   Tetraodon nigroviridis
ORGANISM Tetraodon nigroviridis
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
           Acanthomorpha; Acanthopterygii; Percormorpha; Tetraodontiformes;
           Tetraodontidae; Tetraodontidae; Tetraodon.
REFERENCE 1 (bases 1 to 1643)
AUTHORS   Genoscope.
DIRECT SUBMISSION
TITLE      Submitted (10-AUG-2004) Genoscope - Centre National de Sequencage -
JOURNAL   : 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
COMMENT   (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
           The sequences are based on single pass reads.
           More information available at
           http://www.genoscope.cns.fr/tetraodon.
FEATURES
   source
       1..1643
           /organism="Tetraodon nigroviridis"
           /mol_type="mRNA"
           /db_xref="taxon:99883"
           /tissue_type="fish"

ORIGIN
Query Match      24.5%; Score 386.8; DB 3; Length 1643;
Best Local Similarity 60.9%; Pred. No. 3.4e-93;
Matches 798; Conservative 0; Mismatches 452; Indels 60; Gaps 8;

```

```

Qy      160  ATCATGCAAGTATTCTCGTCGCGAAATTTCTCGACTCCCGGGTAAACCCACCGTCGAGCA 219
Db      113  ATCTTGAAGATCCACGCTCGGAAATATTTCGACTCCCGGGCAACCCACCGTGAGGTTC 172
Qy      220  GAGGTTTTCTCGATGACGGTTCCACCGTGTGCGAGGTGTTCATTCGCGCGCATCCACC 279
Db      173  GACCTGTACACAAAGAAAGGTCTGTTAGAGCTGG---GTCCCAGCGCGCTCCACC 229
Qy      280  GCGCTCCAGAGGCTCATGAGCTGCGTACGCGTGGCGA---TCGCTACTCTGGGCAAGGCG 336
Db      230  GGCATCTACAGGCGCTTGAGCTCCGCGACAAACCAAAACCCGCTACATGCGCAAGGT 289
Qy      337  GTTTTGAAGGCAATTGAACCGTCAACGAGAAATCGGCGACGAGCTCGTGGCGCTAGAG 396
Db      290  GTCTCTAAAGCTGTGAGAAATATCAATAAATTAATTGCACTGCACTGGTTGGCAAGGAC 349
Qy      397  GCTGACGATCAGCGCC-----TCATCGACGAAGCAATGATCAAGCTTGTATGGCACCGCC 450
Db      350  GTCAAGTTCTGGACCGCAAGCAAAATCGACAAGCTGATGCTGGAGATGCGACGCGACCGAC 409
Qy      451  AACAAAGTCCCGCTGGGTGCAACGCAATCTTGGTGTTCATGCTGTTTCCATGGCTGTTGCAAGGCT 510
Db      410  AACAAATCTAAGTTCCGGGCGCAACGCAATCTTGGGCGTGTCCCTGGCTGTGTGCAAGGCT 469
Qy      511  GCTGCTGATTCGCGAGGCTCCCACTGTTCCGCTACATCGGTGGAGCCAAACGC-----563
Db      470  GGTGACGACGAGAAAGGCGTGGCCCTCTACCGCCACATCGCCGACCTGGCGGCAACCC 529
Qy      564  --ACAGTTCTTCCAGTTCCAATGATGAACATCATCAACGGTGGGCGCTCACGCTGACTCC 621
Db      530  CAAGTCATCTCTCCCGTTTCAAGCTCATCAACGGGCGTCCCAACGCGAGCAAC 589
Qy      622  GGTGTTGACGTTACAGAAATTCATGATCGCTTCAATCGGTGCGAGACCTTCTCTGAGGCT 681
Db      590  AAGCTGGCCATGCAAGGATTCATGATCTCTCGCGGTGGAGCCAGCAGCTTCAAAGAGGCG 649
Qy      682  CTCCGCAACGGCGCGAGGCTTACCACGCACTGAAGTCGTCATCAAGGAAAGAGGCG---738
Db      650  ATGCGCATCGGCGCGAAGTCTACCAACACTGAAGACGTCTATCAAGAGAGATGACGCG 709
Qy      739  ---CTGTCCACCGGACTTGGCGATGAGGGCGGCTTCGCTCTTTCCTCGGCTCCACCGCT 795
Db      710  AAGGATGCCACCAACGTTGGGAGACGAGGAGGCTTTTGGCCCAACATCTCTGGAGAAACAAG 769
Qy      796  GAGGCTCTTGACCTTATCGTTAAGGCAATCGAGAAGGCTGGCTTCAACCCAGCAGCAAGAC 855
Db      770  GAAAGCTCTGGAGCTGCTGAAGAACGCCATCGCTAAGGCGGCGCTACAC-----CGACAAG 823
Qy      856  ATCGCTCTTGTCTCTGACGTTGCTTCTCTGAGTTCTTCAAGGACGCGCACCTTACCACTTC 915
Db      824  ATTGTGATCGGATGACGTTGGTGTGCTCCGAGTTCTACAAGGGCGGCAAGTACGACCTG 883
Qy      916  GAA-----GGTGGCGAGCACTTCGCGACGCTCGAGTGGCAAAACGTT 954
Db      884  GACTTCAAGTCTCCGAGACCCCGCGCTACATTTCTCCCGAGCAGCTGGCGGCGCTC 943
Qy      955  TAGCTGAGTCTGTTGACGGTACCCAAATCGTCTCCATCGAGACCCACTGCGAGGAGAT 1014
Db      944  TACAGGGGCTTCGTCAAAGATTACCCAGTGGTGTTCATCGAGACCCCTTTTGACAGGAC 1003
Qy      1015  GACTGGGAGGTTTACACCAACCTCACGCCAAACCATCGGCGACAAGGTTTCAGATCGTTGGC 1074
Db      1004  GACTGGGAGGCTGGACCAATTTCAAGCCAGCACCCAGC-----ATTCAAGTGGTGGCG 1057
Qy      1075  GACGACTTCTTCGTCACCAACCTCTAGCGCTGAAAGAGGCGCATCGCTAAGAGAGGCTGCC 1134
Db      1058  GACGACCTCACGGTCAACCAACCCCAACGCATCGCCAAAGGGTGTGGGCCCAAGTCTCTGC 1117
Qy      1135  AACTCCATCTGTTAAGTGAACCGATCGGTACCTTACCAGACCTTCCGAGCTTTCGAGCTGTC 1194
Db      1118  AACTGCCTGTGCTCAAGTCAACCCAGATCGGCTCGGTCCAGGATCCCTTCGAGGCGATGC 1177

```



```

ACCESSION   BM3211196
VERSION     BM3211196.1  GI:18055602
KEYWORDS    EST
SOURCE      Mastigamoeba balamuthi
ORGANISM    Mastigamoeba balamuthi
REFERENCE   1 (bases 1 to 1505)
AUTHORS     Bapteste, E., Brinkmann, H., Lee, J. A., Moore, D. V., Sensen, C. W.,
            Gordon, P., Durufle, L., Gaasterland, T., Lopez, P., Muller, M. and
            Philippe, H.
TITLE       The analysis of 100 genes supports the grouping of three highly
            divergent amoebae: Dictyostelium, Entamoeba, and Mastigamoeba
JOURNAL     Proc. Natl. Acad. Sci. U.S.A. 99 (3), 1414-1419 (2002)
MEDLINE     21819461
PUBMED      11830664
COMMENT     Contact: Muller Miklos
            Laboratory of Biochemical Parasitology
            The Rockefeller University
            1230 York Avenue, New York, NY 10021, USA
            Email: mmuller@rockvax.rockefeller.edu
            Insert Length: 1505 Std Error: 0.00
            POLYA=Yes.

FEATURES             Location/Qualifiers
     source           1..1505
                     /organism="Mastigamoeba balamuthi"
                     /mol_type="mRNA"
                     /strain="ATCC 30984"
                     /db_xref="taxon:108607"
                     /clone_lib="Mastigamoeba balamuthi lambda ZAP II Library"
                     /notes="syn: Phreatamoeba balamuthi"

ORIGIN
Query Match      24.28; Score 382.6; DB 4; Length 1505;
Best Local Similarity 59.98; Pred. No. 4.6e-92;
Matches 781; Conservative 0; Mismatches 469; Indels 53; Gaps 7;

QY 143 AGGCACAGTGGTGAATCATGCACGTATTTCGCTCGCGAAATTCCTGACTCCCGCGGTA 202
DB 27 AAGCAACAATGTGACCATCAAGAGCGTCTTCGCCGCGAGANCCCTGGACTCTCGTGGCA 86
QY 203 ACCAACCGTCGAGCGAGAGGTTTCTGTGATGACGTTCCACGGTGTGCGAGGTGTC 262
DB 87 ACCCCACCGTCGAGGTGTGACTGACCAACGAGAGGCGCTCTTCGG---CTCGCGCGTCC 143
QY 263 CATCCGCGCGCATCACCGCGCTCCACGAGCTCATGAGCTGCGTGAAGGTCGGA---TC 319
DB 144 CCTCTGGTCTGACTGGCATCTATGAGCGCTGCGAGCTCCGTGACGGCGACAAGTCGC 203
QY 320 GTACTCTGGCAAGGGCGTTTGAAGCAGTTGAAACCGTCAACGAGAAATCGGCGACG 379
DB 204 GCTACTCTCGGCAAGGGTGTCTGAAGCGCTGAGAACGTCGAGAACGATCCTCGCGCCA 263
QY 380 AGCTCGCTGCGCTAGAGGTGACGATCAGCGCTCATCGACGAGCAATGATCAAGCTTG 439
DB 264 AGCTGATCGCGCTCGAGCTCACCAAGCGGCGAGATCGACAGGCTCATGCTCCAGATCG 323
QY 440 ATGCGACCGGCAACAAGTCCCGCTGGTGCACCAAGCAATCTTGTGTTTCCATGGCTG 499
DB 324 ACGGCACTTAGACAAGACCACTGGCGCCCAACGCCATCTTCGGTGTGCTGATGCGG 383
QY 500 TTGCAAGAGGTGCTGTGTGATTCGCGAGCGCTCCACCTGTTTCGCTATACATCG----- 550
DB 384 TCTGCGGTGCGCGCGCGGTTCAGGGCGCTGCGCCCTCTACCGCTACATCGCGGANNGT 443
QY 551 GTGGACCAACGACAGGTTCTTCCAGTTCATGATGAACATCATCAACGGTGGCGTTC 610
DB 444 CGGGCAACAAGTGCCTCATGCTGCTCTTCNNNTTCAACGTCATCAACGGCGGTGAGC 503
QY 611 ACCTGACTCCGGTGTGTGAGTTCAGGAATTCATGATCGCTCCAAATCGGTGACAGACCT 670
DB 504 ACGCGGGCAACAGCTCGCATCAGGAGTTCATGATCTGCCCAACGGGCGACCTCGT 563
QY 671 TCTCTGAGGCTCTCCGCAACGGCGCGAGGTCTTACCACGCACTGAAGTCGTCATCAAGG 730

Db 564 TCACGAGGCCCTGGCATGCGCGCGAGAGCTACCANAACTGAAGCTGTNATCAAGA 623
QY 731 AAAAGGGC-----CTGTCCACCGGACTTGGCGATGA-GGGCGGCTTCGCTTCCTCCGTC 783
DB 624 AGAGGTACGSCATGGACGCGACCAACGTCGCGCAGGAGGGTTCCTTCGCCCCCAACATC 683
QY 784 GGCTCCACCGGTGAGGCTCTTGACCTTATGCTTAAAGCAATCAGAAAGGCTGGCTTCAAC 843
DB 684 CAGGCCAAACACAGGCGCTTCGAGCTCATCGTTGAGGCCATCAAGCAGGCGCGGTACACT 743
QY 844 CCAGGCAAGACATCGCTCTGCTCTGGAGCTTGTCTCTGAGTTCCTTCAAGGACGGC 903
DB 744 GGCAAGATCGAGATCGGATCGGACGTTGCCGCTCTGCTTCTGGGACGCGCAAGAGTCG 803
QY 904 ACCTACACTTCGAAGGTGGCCAGCACTCCGAGCTGAGATGCAACAGTTTACGCTGAG 963
DB 804 AAGTACGACCTCGGCTTCAAGGTCCCGCGCAGCAAGAAGACCCCGACATGCTGNTCG 863
QY 964 CTCGTTGACGC-----GTACCAATCTGCTTCATC 993
DB 864 GCGAGGNCCTTATCAAGCTCTACGAGGATGGACCTCGAAGTACCGGATCTGTCGATC 923
QY 994 GAGGACCCACTCGCAGGAAGATGACTGGAGGTTTACACCAACCTCACCGCAACCATCGGC 1053
DB 924 GAGGACCCCTTCGACCCAGAGAGACTGGGCCACTACACAGGTTTACCGAGCTGATCCGC 983
QY 1054 GACAAGTTTCAGATCGTTGGCGACGACTTCTTCGTCAACCAACCTCAGCGCTCGAAGGAG 1113
DB 984 AACAGGATCCAGATCGTCGCGGACGACCTGCTGTTCAACCAACCCCAAGCGCATCGTCG 1043
QY 1114 GGCTCGCTAAGNAGGCTGCCAATCTCATCTCTGTTAAGTGAACAGATCGGTACCGTTC 1173
DB 1044 GCCCGCAACAAGAAAGCGCTGCAACGCGCTGCTCAAGCTCAACAGATCGGCTCGGTC 1103
QY 1174 ACCGAGACCTTCGACGCTGTGACATGCTCACCGCGGAGGTACACCTCCATGATGTC 1233
DB 1104 AGCGAGCGCTCGAGCGTCCG--CTTGGCGCGGAGGTCAACTGGGGCGTCATGTTGTC 1162
QY 1234 CACCGTTCGGTGAAGACCGAGGACACCAATTCGCTGACCTCGCAGTTGCACTCAACTGT 1293
DB 1163 CACCGTTCGGGCGAGACTGAGGACGCGTTCATCGCGGACCTGCTCGCTCGTGGGTCG 1222
QY 1294 GSCACATCAAGACTGCTGCTCCAGCAGCTTCGACCGGTGTCGAAGTACACACGCTT 1353
DB 1223 GGCAGATCAAGACTGCTGCTCCCTCGAGGCGCTCGCAAGTACACACGCTC 1282
QY 1354 CTCGCGATCGAGCAGCTGCTTGGCGACGCGCGCTCTACGCGAG 1396
DB 1283 CTCGCGATCGAGGAGGAGCTCGGCGCAACGCCACTACGCGG 1325

RESULT 7
LOCUS      CR698859 1613 bp mRNA linear HTC 19-AUG-2004
DEFINITION Tetraodon nigroviridis full-length cDNA.
ACCESSION  CR698859
VERSION    CR698859.1 GI:51196768
KEYWORDS  HTC; cDNA; full-length; Tetraodon nigroviridis.
SOURCE     Tetraodon nigroviridis
ORGANISM   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
            Acanthomorpha; Acanthopterygii; Percormorpha; Tetraodontiformes;
            Tetraodontidae; Tetraodontinae; Tetraodon.
REFERENCE  1 (bases 1 to 1613)
AUTHORS    Genoscope.
TITLE      Direct Submission
JOURNAL    Submitted (10-AUG-2004) Genoscope - Centre National de Sequencage -
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            [E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr]
            The sequences are based on single pass reads.
            More information available at

```


QY 280 GCGTCCACGAGGCTCATGAGCTGCGTGACGGTGGCGA--TCGC

Db 198 GGCATCTACGAGGCCCTGGAGCTCGCGCAACACGACAAAACCCGCTACATGGGCAAGGT 257
Qy 337 GTTTTGAAGGAGGTGAAACGTCACAGAAAGAAATCGCGCAGAGCTCGCTGGGCTAGAG 396
Db 258 GTCTCTAAAGCTGTGAGAAATATCAATAAATAAATATGACCTGCACTGGTTGGCAAGGAC 317
Qy 397 GCTCAGCATCAGCGCC-----TCATCGACGAAGCAATGATCAAGCTTGATGGACCGCC 450
Db 318 GTCAACGTTCTGGACAGGCAAAAATCGAAGCTGATGCTGGAGATGGAGCGGACCGAC 377
Qy 451 AACAAAGTCCCGCTGGGTGCAAAACGCAATCCTGGTGTTCATGCGTGTGTGCAAGGCT 510
Db 378 AACAAATCTAAGTTCGGGGCAACGCCNTCTGGCGTGTCCCTGGTGTGTGCAAGGCT 437
Qy 511 GCTGCTGATTCGCGAGGCTCCCACTGTTCCGCTTACATCGGTGACCAAAACGC----- 563
Db 438 GGTGCAGCAGAAAGGCGGTGCCACTCTACCGCACATCGCGACCTGGCGGCAACCCC 497
Qy 564 --ACACGTTCTCCAGTTCCTCAATGATGAATCATCATCAAGTGGCGCTCAGCTGACTCC 621
Db 498 CAAGTCACTCTCCCGTTCGCGTTTCAACGTCATCAACGCGGCTCCCAAGGAGAAC 557
Qy 622 GGTGTGACGTTTCAGGAATTCATGATCGCTCCAAATCGGTGACAGACCTTCTCTGAGGCT 681
Db 558 AAGCTGGCCATGCAAGGAGTTTATGATCTCGCGTGGGAGCGCAGCAGCTTCAAAGAGGCC 617
Qy 682 CTCGCAACGCGCGAGGTCTTACCAGCACTGAAGTCCGTTCATCAAGAAAGAGGCG--- 738
Db 618 ATGGCATCGGCGCGAGGCTTACCACAACCTGAAGAACGTCATCAAGGAGAAATGACGGC 677
Qy 739 ---CTGTCCACCGGACTTGGCGATGAGGCGGCTTCTGCTCTTCCGTCGGCTCCACCGGT 795
Db 678 AAGGACGCCAACCAACGTTGGAGAGCAGGAGGCTTGTCCCGCCACATCTCTGGAGAACAG 737
Qy 796 GAGGCTCTTGACCTTATCGTTAAGGCAATCGAAGGCTGGCTTCAACCCAGGCAAGGAC 855
Db 738 GAAGCTCTGAGTGTCTGAAGAACCGCATCGCTAAGGCGCGCTACAC-----CGACAAG 791
Qy 856 ATGCTCTTCTCTGAGGTTGCTTCTCTGAGTCTTCAAGACGCGCAGCTTACCACTTC 915
Db 792 ATTGTGATCGGATGAGGCTGCTCTGCTCTGAGTCTTCAAGGCGGCGCAAGTACGACCTG 851
Qy 916 GAA-----GGTGGCCAGCAGCTCCCGCAGCTGAGATGGCAACAGTT 954
Db 852 GACTTCAAGTCTCCGAGACGCCCGCGCTACATCTCTCCGAGCAGCTGGCGGACCTC 911
Qy 955 TAGCTGAGTCTGTTAGCGGTACCCATCGTCTCCATCGAGACCCACTGCAAGGAGAT 1014
Db 912 TACAGGGGCTTCGTCAAAGATTACCCAGTGTGTTCATCGAGGACCCCTTTGACAGGAC 971
Qy 1015 GACTGGAGGTTTACACCAACCTCACCGCAACCATCGGCGACAAGTTTCAGATCGTTGGC 1074
Db 972 GACTGGAGGCTGGACCAATTTCAACAGCAGCAGCAGC-----ATTGAGGTGGTGGGC 1025
Qy 1075 GACGACTCTTCTGTCACCAA--CCCTGAGCGCTTGAAGGAGGCGATCGCTAAGAGGCTGC 1133
Db 1026 GACGACCTCAGCTCACCAACCCCAACGATCGCCAGAGGTTGCGCCAGAGTCTCTG 1085
Qy 1134 CAACCTCATCTCTGGTTAAGGTGAACAGATCGGTACCTCACCAGACCTTCAAGCCTGT 1193
Db 1086 CAACCTGCTGCTCTCAAGTCAACAGATCGGCTCGGTACCGGCTCGCTTCAAGGAGCATG 1145
Qy 1194 CGACATGGCTCACCGGCGAGGCTACACCTCCATGATGTCCACCGTTCCGCTGAGACCGA 1253
Db 1146 CAAGATGGCCAGAGCAACGCTGGGCGGTGATGGTCAGGCATCGCTTCGGGGAGACGGA 1205
Qy 1254 GGAACACCAATTGCTGACCTCGAGTTGCACTCAAATGCTGGCCAGATCAAGACTGGTGC 1313
Db 1206 GGACACCTTTCATCGCTGACCTGGTGTGCTGTGTGTGTCACCGGACAGATCAAGACGGCGC 1265
Qy 1314 TCCAGCACGTTCCGACCGGTGTCGAAGTACAAACAGCTTCTCCGATCGAGCAGTGTCT 1373
Db 1266 TCCGTGCCGCTCCGAGGCTTGGCCAAATACAAACAGCTGCTCCGGATCGAGGAGCT 1325

Qy 1374 TGGCGACGCGCGGCTCTAAGCAGTCCGACGCGATTCACACGCTTTTCAGG 1424
Db 1326 GGCGCAGCAGGCGGTCGCGCGAAGAACTTCAGGACCCCATCTGAGCG 1376

RESULT 9
CR728153 1637 bp mRNA linear HTC 19-AUG-2004
LOCUS Tetraodon nigroviridis full-length cDNA.
DEFINITION
ACCESSION CR728153
VERSION CR728153.1 GI:51226414
KEYWORDS HTC; cDNA; full-length; Tetraodon nigroviridis.
SOURCE Tetraodon nigroviridis
ORGANISM Tetraodon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodontidae; Tetraodon.
REFERENCE 1 (bases 1 to 1637)
Genoscope.
AUTHORS Direct Submission
TITLE Submitted (10-AUG-2004) Genoscope - Centre National de Sequencage -
JOURNAL : 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
COMMENT : E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
The sequences are based on single pass reads.
More information available at
http://www.genoscope.cns.fr/tetraodon.

FEATURES
Location/Qualifiers
1..1637
/organism="Tetraodon nigroviridis"
/mol_type="mRNA"
/db_xref="taxon:99983"
/tissue_type="fish"

ORIGIN
Query Match 24.0%; Score 378.6; DB 3; Length 1637;
Best Local Similarity 60.5%; Pred. No. 5.7e-91;
Matches 792; Conservative 0; Mismatches 459; Indels 58; Gaps 8;
Qy 160 ATCATGACGATTTCTGCTCGGAAATTCCTGACTCCCGGGTAAACCAACCGTCGAGGCA 219
Db 108 ATCTTGAAGATCCACGCTCGGAAATATTCGACTCCCGGGCAACCCACCGTGGAGGTC 167
Qy 220 GAGGTTTCTCGATGACGGTTCACCGGTTCGAGGTGTTCCATTCGCGGCATCCACC 279
Db 168 GACCTGTACCAAGAAGGCTGTACAGAGCTGCGGTCC--CCAGCGCGCTCCACC 225
Qy 280 GCGTCCACGAGGCTCATGAGCT--GCGTGACGGTGGCGATCGCTACCTGGGCAAGGGCG 337
Db 226 GGCATCTAGAGGCTTGGAGCTCCGCAACGACAAACACCGCTACATGGGCAAGGTG 285
Qy 338 TTTTGAAGGAGTTGAAACAGTCAACGAAGAAATCGGCGACGAGCTCGCTGCGCTAGAGG 397
Db 286 TCTCTAAAGCTCTGAGAAATATCAATAAATAATTTGCACCTGCATCGTTGGCAAGGACG 345
Qy 398 CTGACGATCAGCGCC-----TCATCGAGGAAGCAATGATCAAGCTTGATGCAACGCCA 451
Db 346 TCAACGTTCTGGACAGGCAAAAAATCGACAAAGCTGATGCTGGAGATGGACGCAACGCA 405
Qy 452 ACAAGTCCCGCTCGGTGCAAAACGCAATCCTTGGTGTTCATGGCTGTGTCAAAGGCTG 511
Db 406 ACAATCTAAGTACGGGGCAACGCCATCTCGGCGTGTCCCTGCTGTGTGTCGAAGGCTG 465
Qy 512 CTGCTGATTTCCGAGGCTTCCACCTGTTCCGCTACATCGGTGGAACAAACGC----- 563
Db 466 GTGCAGCAGAGAGGCGGTGCCCTCTACCGCCACATCGCGACCTGGCGCGCAACCCCC 525
Qy 564 -ACAGCTTCTCCAGTTCCTCAATGATGAACATCATCAAGTGGCGCTCAGCTGACTCG 622
Db 526 AAGTCATCTCTCCCGTTCCTCGCTTTCAACGCTTCAACGCGCGCTCCCAAGGAGCAACA 585
Qy 623 GTGTTGACGTTTCAGGAATTCATGATCGCTCCAAATCGGTGCGAGAGACCTTCTCTGAGGCTC 682

```
Db 586 AGCTGGCCATGCGAGGAGTTTCATGATCTGCGGTGGAGCCAGCAGCTTCAAGAGGCCA 645
Qy 683 TCCTCAACCGCGGGAGGCTTACACGCACTGAAGTCCGTATCAAGGAAAGGCG--- 738
Db 646 TGGCATCGCGCGGAGTCTACCAACAACCTGAAGACGTCATCAAGGAGAGTACGGCA 705
Qy 739 --CTGTCCACCGGACTTGGCGATGAGGGCGGCTTCGCTCTTCGTCGGTCCACCGGTG 796
Db 706 AGGACGCCACCAACGTTGGGAGACGAGGGAGGCTTTGCCCCCAACATCTGGAGAACAGG 765
Qy 797 AGGCTCTTGACCTTATCGTTAAGGCATCGAGAGGCTGGCTTCAACCCAGGCAAGACA 856
Db 766 AAGCTCTGGAGCTGCTGAAGAACGCCATCGCTAAGGCGGCTACAC-----CGACAGA 819
Qy 857 TCCTCTTGCTCTGGAGGCTTCTCTCTCTGAGTCTTCAAGGACGGCAGCTTACACCTTCG 916
Db 820 TTGTGATCGCATGGAGCTGGCTCGCTCCGAGTTCTTCAAGGGGGGCAAGTACACCTGG 879
Qy 917 AA-----GGTGGCCAGCACCTCCGAGCTGAGATGGCAAAAGTTT 955
Db 880 ACTTCAAGTCTCCCGACGACCCCGCGCTACATTTCTCCGAGCAGCTGGCCGACCTCT 939
Qy 956 ACCTGAGCTCTGAGCGCTACCCATCGTCTCCATCGAGGACCCACTCGCAGAGATG 1015
Db 940 ACAGGGGCTTCTGTCAAAGATTACCCAGTGTGTCATCGAGGACCCCTTTGACCGAGC 999
Qy 1016 ACTGGGAGGTTACACCAACTCACCGCAACCATCGCGCACAGGTTTCAGATCGTTGGCG 1075
Db 1000 ACTGGGAGGCTGACCAATTTTACAGCAGCACAGC-----ATTTCAGGTGGTGGCG 1053
Qy 1076 ACAGCTTCTTCTGACCAACCTGAGCGCTGAGGAGGAGCTGCTTAAGAGGCTGCCA 1135
Db 1054 ACAGCTTACGGTCAACCAACCCCAACGATCGCAAGGCTGTGGCCAGAGTCTCTGCA 1113
Qy 1136 ACTCCATCTGTTAAGGTGAACAGATCGGTACCTTCACCGAGACCTTCGACGCTGTG 1195
Db 1114 ACTGCTCTGCTGCTCAAGTCAACAGATCGGCTCGGTCAACGAGTCTCCGAGCATGCA 1173
Qy 1196 ACATGCTCAACCGCGAGGCTACACCTCCATGATGTCACACCTTCGGTGAGACCGAG 1255
Db 1174 AGATGCCCAAGACAACGGCTGGGGGTGATGTCAGCCATCGTCCGGGAGACGGAG 1233
Qy 1256 ACACCACTTGTGACCTCGCAGTTGCACTCAACTGTGCGCCAGATCAAGACTGGTGCTC 1315
Db 1234 ACACCTTCACTGCTGACCTGCTGCTGCTGTCGACCGGACAGATCAAGACGGCGCTC 1293
Qy 1316 CAGCAGTTCGACCGGTGTCGCAAGTACAAACAGCTTCTCCGATCGAGCAGCTGCTTG 1375
Db 1294 CGTGGCGCTCCGAGCGCTTGGCCAAAGTACAAACAGCTGCTCCGGATCGAGGAGAGCTG 1353
Qy 1376 GCGACCGCGCTGTCAGGAGTTCGAGCGCATTCACACCTTCAGGG 1424
Db 1354 GCGACCGCGCTTCCGCGGCAAGAACTTCAGGCAACCCCATCTGAGCG 1402
```

RESULT 10

```
CR685932
LOCUS Tetraodon nigroviridis full-length cDNA.
DEFINITION
ACCESSION CR685932
VERSION CR685932.1 GI:51183839
KEYWORDS HTC; cDNA; full-length; Tetraodon nigroviridis.
SOURCE Tetraodon nigroviridis
ORGANISM Tetraodon nigroviridis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
            Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
            Tetraodontidae; Tetraodontidae; Tetraodon.
REFERENCE 1 (bases 1 to 1551)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (10-AUG-2004) Genoscope - Centre National de Sequencage -
```

```
1 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
(E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
The sequences are based on single pass reads.
More information available at
http://www.genoscope.cns.fr/tetraodon.
Location/Qualifiers
source 1..1551
        /organism="Tetraodon nigroviridis"
        /mol_type="mRNA"
        /db_xref="taxon:99883"
        /tissue_type="Eyes"

COMMENT
Query Match 23.5%; Score 371.2; DB 3; Length 1551;
Best Local Similarity 60.4%; Pred. No. 5.7e-89;
Matches 783; Conservative 0; Mismatches 453; Indels 60; Gaps 8;
```

```
ORIGIN
174 CGTCCGGAATTCCTGACATCCCGCGGTAAACCAACCGTCGAGCAGAGGTTTCTCGGA 233
27 CGTCCGGAATTCCTGACATCCCGCGCAACCCACCGTGGAGTCAACCTGTACACCA 86
234 TGACGGTTCACCGGTGTCGAGTGTCCATCGCGGCATCCACCGCGTCCACGAGGC 293
87 GAAGGTCTGTTTTCAGAGCTGCG---GTCCCCAGCGCGCTTCCACCGCATCTACGAGC 143
294 TCATGAGCTGCGTG---ACGGTGGCGATCGCTACTCTGGGCAAGGGCGTTTGAAGGCGT 350
144 CTTGGAGCTCCCGACACAGATTAAACCGCTACATGGGCAAGGGGTCAAAAGGCGT 203
351 TGAACAGCTCAACAGAAATTCGGGACAGCTCGTGGCTTAGAGGCTGACGATACGCG 410
204 TAAATATATCAATGAGTTTTTGGCCCCCGCTTGTGTAAACGAGACGCTCAACGTTCTGGA 263
411 CC-----TCATCGAGCAAGCAATGATCAAGCTTGTATGGCACCGCCCAACGATCCCGCT 464
264 CCAGGCAAAATCGACAAGCTGATGCTGGAGATGGACGGCACCGACCAACAAATCTAAGTT 323
465 GGGTCAAAACGCAATCTTGGTGTTCATGGTGTTCGCAAGGCTGCTGCTGATTCGCG 524
324 CGGGCCACAGCCATCTTGGCGTGTCTTGGCTGTGTGCAAGGCTGGTGCAGCAGAGAA 383
525 AGGCTCCCACTGTTCCGTATACATCGGTGGACCAAAAG-----CACACGTTCTTCC 575
384 GGGCGTGGCTCTTACCGCCACATTCGCGACCTGGCGGCAACCCCAAGTCATCTCTCC 443
576 AGTTCAATGATGAACATCATCAACGGTGGCTCAGCTGACTCCGGTGTGACGTTC 635
444 CGTTCCCGCTTTTCAACGTTGATCAACGGCGCTCCACGAGGCAACAAAGCTGGCCATGCA 503
636 GGAATTCATGATCGCTCCAATCGGTGCAGAGACCTTCTCTGAGGCTCTCCGCAACGGCG 695
504 GGAGTTTATGATCTCGCGGTGGAGCCAGCAGCTTCAAGAGGCGCATGCGATCGGCGC 563
696 GGAGGTCTATACCGCAGCTGAAGTCCGTCATCAAGGAAAGGGG-----CTGTCCACCGG 749
564 CGAGGTCTACCAACACCTGAAGAACGTCATCAAGGAGAGTACGGCAAGGACGCCACCA 623
750 ACTTGGCGATGAGGGCGCTTCTGCTCTTCCGTGGCTCCACCGGTGAGGCTCTTGACCT 809
624 CGTGGGAGACGAGGAGGCTTTGCCCGCCCAACATCTCTGGAGAACAAAGAAAGCTCTGAGCT 683
810 TATCGTTAAGGCAATCGAGAGGCTGGCTTCAACCCAGGCAAGGACATCGCTCTGCTCT 869
684 GCTGAAGAACGCCATCGTAAGCGCGGTACAC-----CGACAAGATTGTGATCGGCAT 737
870 GGACGTTGCTCTCTGAGTTCTTCAAGGACGGCAGCTTACCACTTCGAA----- 918
738 GGACGTTGCTCTCTGAGTTCTTCAAGGCGGCAAGTACGACCTGGACTTCAAGTCCCC 797
919 -----GGTGGCCAGCAGCTCCGAGCTGAGATGGCAAAAGTTTACGCTGAGCTCGT 968
798 CGACGACCCCGCGCTGATCATCTCTCCGAGCAGCTGGCGGACCTCTACAGGGGCTTCGT 857
```


QY 1313 CTCAGCAGCTTCGACCGTGTGCGAAAGTACACAGCTTCTCCGATCGAGCAGCTG 1372
 |||||
 Db 1293 CTCGTCGCGCTCCGAGCGCTTGGCCAAAGTACACAGCTGCTCCGATCGAGAGAGC 1352
 |||||
 QY 1373 TTGGCGAGCGCGGCTTACGAGGTGCGAGCGGATTTCCACGCTTTCAGGG 1424
 |||||
 Db 1353 TGGCGAGCAGCGCCCTTCGCGCGAAGAACTTCAGGCACCCATCTGAGCG 1404
 |||||

RESULT 12
 CR656482
 LOCUS
 DEFINITION Tetraodon nigroviridis full-length cDNA.
 ACCESSION CR656482
 VERSION GI:51152927
 KEYWORDS HTC; cDNA; full-length; Tetraodon nigroviridis.
 SOURCE Tetraodon nigroviridis
 ORGANISM Tetraodon nigroviridis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
 Tetraodontidae; Tetraodontidae; Tetraodon.
 1 (bases 1 to 1582)
 Genoscope.
 Direct Submission
 Submitted (10-AUG-2004) Genoscope - Centre National de Sequencage -
 2 rue Gaston Cremieux, CP 5705 - 91057 EVRY cedex - FRANCE
 (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
 The sequences are based on single pass reads.
 More information available at
 http://www.genoscope.cns.fr/tetraodon.

FEATURES
 Location/Qualifiers
 1..1582
 /organism="Tetraodon nigroviridis"
 /mol_type="mRNA"
 /db_xref="taxon:99883"
 /tissue_type="Liver"

ORIGIN
 Query Match 23.2%; Score 366.2; DB 3; Length 1582;
 Best Local Similarity 60.4%; Pred. No. 1.3e-87;
 Matches 792; Conservative 0; Mismatches 458; Indels 61; Gaps 9;

QY 160 ATCATGCAGCTATTCGCTCCGAAATTCGACTCCCGGGTAACCAACCGTCGAGGCA 219
 |||||
 Db 67 ATCTGAAGTCCAGCTCGGAAATATTCGATCCCGCGCAACCCACCGTGGAGTC 126
 |||||
 QY 220 GAGTTTCTCGATGACGTTCCACGTTGTCGAGGTGTTCCATCCGCGCATCCACC 279
 |||||
 Db 127 GACCTGTACACCAAGAAAGCTGTTTCAGAGTGGC---GTCCCGAGCGCGCTCCACC 183
 |||||
 QY 280 GGGTTCACAGAGCTCATGAGCTGCGTAGCGTGGGA---TGGCTACCTGGGCAAGGC 336
 |||||
 Db 184 GGCATCTACAGGCGCTGGAGCTCCGCGCAACGACAAACCCGCTACATGGGCAAGGT 243
 |||||
 QY 337 GTTTTGAAGCAGTTGAACAGTCAACAGAGAAATCGGCACAGCTGCTGCTGCTAGAG 396
 |||||
 Db 244 GTCTTAAGCTGTGAGATATCAATAAATAATTCGACCTGCATGCTGGTTGGCAAGGAC 303
 |||||
 QY 397 GCTGAGATCAGCGCC-----TCATCGACGAAGCAATGATCAAGCTTGTATGGCAGCGC 450
 |||||
 Db 304 GTCAAGCTTCTGACCGAGGCAAAATCGAAGCTGATGCTGGAGATGAGCGCACCGAC 363
 |||||
 QY 451 AACAAAGTCCCGCTGGGTGCAACCGAATCTTGGTGTTCATGCTGTTGCAAAAGGCT 510
 |||||
 Db 364 AACAAATCTAAAAACGGGGCAACGCCATCTCGGGGTGTCTCCCTGTGTGTGCAAGGCT 423
 |||||
 QY 511 GCTGCTGATTTCCGAGCGCTCCACACTGCTCGGTACATCGGTGGAACCAAGC----- 563
 |||||
 Db 424 GGTGAGCAGAGAAAGGCGTGGCACTCTACCGCACATCCGACCTGGCGCGCAACCC 483
 |||||
 QY 564 --ACAGTTCCTCCAGTTCCAATGATGAACATCATCAACCGTGGCGCTCAGCTGACTCC 621
 |||||

Db 484 CAAGTCATCTCCCGCTTCCGCTTTCAAGCTCATCAACGCGGCTCCACGAGGAAC 543
 |||||
 QY 622 GGTGTTGAGCTTCAGGAATTCATGATCGCTCAATCGGTGCAGAGACTTCTCTGAGGCT 681
 |||||
 Db 544 AGCTGCGCATGACGAGGTTTCATGATCTCTGCGGTGGAGCCAGCAGCTTCAAAGAGGCC 603
 |||||
 QY 682 CTCGCAACGCGCGGAGGTCTACACGCACTGAAGTCCGCTCATCAAGGAAGAAAGGGC--- 738
 |||||
 Db 604 ATCGCATCGGCGCGAAGTCTACACACCTGAAGAACGTCTCAAGGAGAGTACGGC 663
 |||||
 QY 739 ---CTGTCCACCGGACTTGGCGATGAGGGGCTTCCGCTCTCTCCGTCGCTCCACCGT 795
 |||||
 Db 664 AAGGACGCCACCAACGCTGGGAGAGGAGGCTTGGCCCAACATCTCTGGAGAACAA 723
 |||||
 QY 796 GAGGCTCTGACCTTATCGTTAAGGCAATCGAAGAGCTGGCTTCAACCCAGGCAAGGAC 855
 |||||
 Db 724 GAAGCCCTGGAGCTGCTGAAGAACGCTATCGTAAGCGGCTACAC-----CGACAAG 777
 |||||
 QY 856 ATCGCTTTGCTCTGGACGTTGCTTCTCTGAGTCTTCAAGGACGCGACCTACCACTTC 915
 |||||
 Db 778 ATTGTGATCGGCATGAGCTGGCTGCTCTGAGTTCTCAAGGGCGGCAAGTACGACTG 837
 |||||
 QY 916 GAA-----GGTGGCCAGCACTCCGCAAGTTCGAGTGGCAAAAGCTT 954
 |||||
 Db 838 GACTTCAAGTCTCCGACGACCCCGCGCTACATCTCTCCGAGAGCTGGCGGACCTC 897
 |||||
 QY 955 TAGCTGAGCTCGTTGACGGGTACCAATCGTCTCCATCGAGGACCCACTGCAGGAAGAT 1014
 |||||
 Db 898 TACAGGGGCTTCGTCAAAGATTACCCAGTGGTGTCCATCGAGGACCCCTTTGACCAAGGAC 957
 |||||
 QY 1015 GACTGGAGGGTTTACACCAACCTCACCGCAACCATCGGCGACAAGGTTTCAGATCGTTGGC 1074
 |||||
 Db 958 GACTGGAGGGGTGGACCAATTTACAGCAGCAGCACAGC-----ATTGAGTGGTGGC 1011
 |||||
 QY 1075 GACGACTTCTTGTGTACCAA--CCCTGAGCGCTTGAAGGAGGAGTTCGTAAGAGGCTGC 1133
 |||||
 Db 1012 GAACGCTCAGGCTCACCAACCCCAACGCAATCGCAAGGCTGTGSCCCAGAAAGTCTG 1071
 |||||
 QY 1134 CAACTCCATCTGTTAAGTGAACGATCGGTACCTCCCTCACCGAGACTTTCAGCTGT 1193
 |||||
 Db 1072 CAATCTGCTGCTCTCAAAAGTCAACAGATCGGCTCGGTACCGAGTCCCTTCGAGGATG 1131
 |||||
 QY 1194 CGACATGGCTCACCGCGCAGGTACACCTCCATGATGTCCACCGCTTCCGGTGCAGACCGA 1253
 |||||
 Db 1132 CAGATGGCCAGAGCAACGGCTGGGGGTGATGGTCAGCCATCGCTCCGGGAGAGCGA 1191
 |||||
 QY 1254 GGACACCACTTGTGACCTCGCAGTTGCATCTCAACTGTGGCCAGATCAAGACTGGTGC 1313
 |||||
 Db 1192 GGACACCTTTCATGTTCTCTGTTGGTGGTCTGTGCACCGGACAGATCAAGACGGCGC 1251
 |||||
 QY 1314 TCCAGCAGCTTCGACCGGTGTGCAAGGTACACAGCTTCTCCGATCGAGCAGCTGCT 1373
 |||||
 Db 1252 TCCGTGCGCTTCGAGCGCTTGGCCAAAGTACAAACAGCTGCTCCGATCGAGGAAGCT 1311
 |||||
 QY 1374 TGGCGAGCGCGGCTCTACGAGTTCGAGCGCATTCCTCCACGCTTTCAGG 1424
 |||||
 Db 1312 GGGCGACGAGCGCGTTCGCGCAAGAACTTCAGGACCCCATCTGAGCG 1362
 |||||

RESULT 13
 BH770828
 LOCUS
 DEFINITION LHMtag569 MG1363 Random Sequence Tag Library Lactococcus lactis subsp. cremoris genomic, genomic survey sequence.
 ACCESSION BH770828
 VERSION BH770828.1 GI:20373785
 KEYWORDS GSS.
 SOURCE Lactococcus lactis subsp. cremoris
 ORGANISM Lactococcus lactis subsp. cremoris
 Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
 1 (bases 1 to 2021)

AUTHORS TITLE JOURNAL COMMENT

Bolotin, A., Ehrlich, S.D. and Sorokin, A.
Studies of genomes of dairy bacteria Lactococcus lactis
Sci. Aliments (2002) In press
Contact: Sorokin A
Genetique Microbienne
INRA

CRJ INRA, Domaine de Vilvert, 78352 Jouy en Josas cedex, France
Tel: 33 1 34 65 25 16
Fax: 33 1 34 65 25 21
Email: sorokine@jouy.inra.fr
best homologue in strain IL1403 is enoB (95%)
Class: shotgun
High quality sequence start: 30
High quality sequence stop: 1993.

FEATURES source

Location/Qualifiers
1..2021
/organism="Lactococcus lactis subsp. cremoris"
/mol_type="genomic DNA"
/strain="MGL1363"
/sub_species="cremoris"
/db_xref="taxon:1359"
/clone_lib="MGL1363 Random Sequence Tag Library"
/note="vector: pSGM2; Site 1: SmaI; Library of
chromosomal fragments of L.lactis strain MGL1363 was
prepared by partial AluI digestion or by sonication."

ORIGIN

Query Match 21.7%; Score 341.8; DB 8; Length 2021;
Best Local Similarity 56.9%; Pred. No. 5.5e-81;
Matches 691; Conservative 0; Mismatches 512; Indels 12; Gaps 3;

QY 159 AATCATGACGATTTTCCTCGCGAAATTCCTCGACTCCCGCGGTAACCCAAACCGTCGAGGC 218
DB |||||
QY 709 AATCGAAACATTCACCGAAGAAATTTTGATTCAGAGGAATCCAAACCGTTGAGGT 768
DB |||||
QY 219 AGAGGTTTTCCTGATGACCGGTTCCACAGGTGTCAGGTGTTTCCATCCGCGCATCCAC 278
DB |||||
QY 769 AGATGTTAGATTGACAGATGGAATTTTGGACGCGGCGGAGTTCCTCAGGAGCATCAAC 828
QY 279 CGGCTCCACGAGGCTCATGAGTCGTCGAGCGTGGCGATCGCTACCTGGCGAAGGCGT 338
DB |||||
QY 829 AGAGATCGTGAAGCGGTGAACTTAAGATGCGGAGCGGACTTCAAGGAAAGGCGT 888
QY 339 TTGAAGGCGAGTTGAAACGTCACGAAAGAAATTCGCGACGAGCTCGCTGGCCTAGAGGC 398
DB |||||
QY 889 TTCTAAGCTGTTGCAAAATGTTATGTTGAAATTTATGAAGCCTTAAAGGCCAATCACC 948
QY 399 TGAGGATCAGGCGCTCATGACGAAGCAATGATCAAGCTTGATGGCACCGCCCAAGTC 458
DB |||||
QY 949 ATTCAATCAAGCAAAATTAACCATCTGATGATTGAACCTTGATGGGACAAAAATAAGTC 1008
QY 459 CCGCTGGGTGCAAAAGCAATCTTGGTGTTCATGGCTGTTGCAAGGCTGCTGCTGA 518
DB |||||
QY 1009 TCCTTAGGTGCCAATATATGGTGTTCATGGCTATTTGCTGTCGCGCAGCTAA 1068
QY 519 TTCGCGAGGCTCCCACTGTTCCGCTACATCGGTGGACCAACGACACGTTCTTCCAGT 578
DB |||||
QY 1069 CAGTGAAGAAATCCCACTTTATCGTTATCTTGTGGAGTAGATTGGAACCTTCTCAA 1126
QY 579 TCCAATGATGAACATCATCAACGGTGGCGCTACCGTGACTCCGGTGTGACGTTCCAGGA 638
DB |||||
QY 1127 -CCATTTTCAATGTATCAATGTTGAGTGCACGCGGATTCAGGAATTTGACGTTCAAGA 1185
QY 639 ATTTCATGATCGCTCCCAATCGGTGCAGAGACTTCTCTGAGGCTCTCCGCAACGGCGCGGA 698
DB |||||
QY 1186 ATTCTTGATTACCCCTGTCAACGCTGCTAGTTTTCGGATGTTTAGAAAAAATCCCAA 1245
QY 699 GGTCTACCGACTGAAGTCCGTCATCAAGGAAAAAGGCGCTGTCCACCGGACTTGGCGGA 758
DB |||||
QY 1246 TATCTATCATCTTGAAGAAAAATTTTAGCTGATAAAGGACTAGAAACGCGCGTAGGGGA 1305
QY 759 TGAGGCGGCTTCGCTTCCTCGCTCCCGCTCCAGGCTTCTTGACCTTATCGTTAA 818
DB |||||

Db 1306 CGAGGGTGGTTTTTCTCCAAAACTTGGTTTCGACGGAATAATGCCAATGCTTACACTTTATCA 1365
QY |||||
QY 819 GGCAATCGAGAGGCTGGCTTACACCCAGCAAGGACATCGCTCTTGTCTCGACGTTGC 878
Db |||||
Db 1366 AGCAATTGAAAGAGCTGGTTATGTCGCCAGGCGAAGAAATTCGATTTGTAATTCACCCGGC 1425
QY |||||
QY 879 TTCTCTGAGTTCCTT-----CAAGGACGGCACCTACCACTTCGAAGGTGGCCAGCACTC 932
Db |||||
Db 1426 TTCTAGCGAATTTTATGATGACAAAGAAAAAGTCTATCATTTTGAAGGTCAAAAGCTCAC 1485
QY |||||
QY 933 CGCAGCTGAGATCGCAAAAGTTTACGCTGAGCTCGTTGACGCGTACCCCAATCG---TCTC 989
Db |||||
Db 1486 TTCTAGCGAATTTTATGACCTAATATGAGGGTTAGTTGAAAAATATCTCGCCCTAATCTC 1545
QY |||||
QY 990 CATCGAGGCCACTGACGAGAAAGTACCTGGGAGGTTTACCAACCACTCACCCGCAACCAT 1049
Db |||||
Db 1546 TATTGAAGACGGATTTTCGGAACATGATTGGGCTGGTTTTCGGCTCAACAGAAAGTTCA 1605
QY |||||
QY 1050 CGCGGACAAAGTTTCAGATCGTTGGCAGCAGCTTCTTCGTCACCAACCTGAGGCGCTGAA 1109
Db |||||
Db 1606 AGGTCAAAAAATTTCAATTTGGTGTGATGATATTTTGTACAAAATCCCGAAATCTTCAA 1665
QY |||||
QY 1110 GGAGGCGATCGCTAAGAAGCTGCCAACTCCATCTCTGGTTAAGGTGAACAGATCGGTAC 1169
Db |||||
Db 1666 AGNAGAAATTCAAAAAGTGTAGCAATGCTATTTTGATTTAACTAAATCAAAATCCGGAC 1725
QY |||||
QY 1170 CTTACCGGAGACTTTCGACGCTGTGCATGCTTCCCGCGAGGCTACACCTTCCATGAT 1229
Db |||||
Db 1726 AGTGACCGAAGCGATTGAAGCGATTAGTTTGGCCAGAAAAAGCAGGTTATAAGACCATGAT 1785
QY |||||
QY 1230 GTCCACCGCTTCGGTGGAGACCGGACACCACTTGTGACCTCGAGTTCGAGTTCACCTCA 1289
Db |||||
Db 1786 TTTCTATCTGTCAGGTGAAACCGTTGATAGTTTATTTGCTGACTTTTGTATAGCCATGCA 1845
QY |||||
QY 1290 CTGTGCGCAGATCAAGACTGGTGTCCAGCAGTTCGACCGTGTCCGAAAGTACAAACA 1349
Db |||||
Db 1846 CGCTGACAAATAAAAAACAGGTTCAATGCGAGAGTGAAGCGGTTGAAAAATATATCA 1905
QY |||||
Db 1350 GCTTCTCCCATCGA 1364
Db 1906 ATTTTGGCTATTGA 1920

RESULT 14 CR733357

LOCUS CR733357 1407 bp mRNA linear HTC 19-AUG-2004
DEFINITION Tetraodon nigroviridis full-length cDNA.
ACCESSION CR733357
VERSION CR733357.1 GI:51231674
KEYWORDS HTC; cDNA; full-length; Tetraodon nigroviridis.
SOURCE Tetraodon nigroviridis
ORGANISM Tetraodon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodontidae; Tetraodon.
1 (bases 1 to 1407)
Genoscope.

REFERENCE

AUTHORS Direct Submission
TITLE Submitted (10-AUG-2004) Genoscope - Centre National de Sequencage -
JOURNAL : 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
(E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)

COMMENT

The sequences are based on single pass reads.
More information are available at
http://www.genoscope.cns.fr/tetraodon.

FEATURES

source
1..1407
/organism="Tetraodon nigroviridis"
/mol_type="mRNA"
/db_xref="taxon:99883"
/tissue_type="fish"

ORIGIN

```

Query Match      20.8%; Score 328.6; DB 3; Length 1407;
Best Local Similarity 59.8%; Pred. No. 1.9e-77;
Matches 687; Conservative 0; Mismatches 40; Indels 53; Gaps 6;

QY 317 ATCGCTACCTCGGCAAGGCGTTTGAAGGAGTTGAAACGTCACGA-----AGAAAT 371
DB 13 ATGCTACTTGGGCAAGGGTCAAAGGGCGTTAAATATGTCAATAGTTTTTGGGCC 72
QY 372 CGGCGAGAGCTCGTGGCTTAGAGGCTGACGATCAGCGCTCATCGACGAAGCAATGAT 431
DB 73 CGCTTGTGTAAACAGGAGCTCAAGTTCTGACACGCAAAATCGCAAGCTGATGCT 132
QY 432 CAAGCTTGATGGCAACCGCAACAGTCCCGCTGGGTGCAAAACGCAATCTTGTGTTTC 491
DB 133 GGAGATGACGGCACCGCAACAAATTAAGTTTCGGGGCCAAACGCCATCTCTGGCGGTG 192
QY 492 CATGGCTTTGCAAGGCTGTGCTGATTCGCGAGGCTCCACATGTTTCGCTACATCGG 551
DB 193 CTTGGCTGTGTGAAGGCTGTGTGACGAGAAAGGGCGTGCCTCTACCGCCACATCGC 252
QY 552 TGGACCAAAACG-----CACACGTTTCTTCCAGTTTCCATGATGAACATCATCAACGG 602
DB 253 CGACTTGGCGGCAACCCCAAGTCATCTCCCGTTCGCGTTTCAACGTCATCAACGG 312
QY 603 TGGCGCTCACGCTGACTCCGGTGTGAGTTTCAGAAATTCATGATCGCTCCAAATCGGTGC 662
DB 313 CGGCTCCACGACGCAACAAAGCTGGCCATGAGGAGTTTCATGATCTCTGCGGTGGAGC 372
QY 663 AGAGACTTCTCTGAGGCTCTCCGCAACGGCGGAGGTCTACCAACGACTGAAGTGGT 722
DB 373 CAGCAGCTTCAAGAGGCGCATGCGCATCGGTGCGGAAGTCTTACCAACCTTGAAGAAGT 432
QY 723 CATCAAGGAAAGGCG-----CTGTCCACGCGACTTGGCGATGAGGCGGCTTCGCTCC 776
DB 433 CATCAAGGAGAGTACGCGAAGGACGACCAACAGTGGGAGACGAGGAGGCTTTGCCCC 492
QY 777 TTCGCTGGCTCCACCGTGAGGCTCTTGACTTATCGTTAAGCAATCGAAGAGGTGG 836
DB 493 CAACATCTGGAGAACAGGAAGCTCTGGAGCTCTGAAGAACGCCATCGTAAGGCGG 552
QY 837 CTTACCCCGGAGGACATGCTCTGCTGCGAGTGTGCTTCTCTGAGTTCTTCAAA 896
DB 553 CTACAC-----CGACAAGATTTGATCGGATGAGCTGGCTGCTCCGAGTTTCTACAA 606
QY 897 GGAGCGCACTTACCACTTCGA-----AGTGGCGGACGACTCCG 935
DB 607 GGGCGGCAAGTACGACTGACTTCAAGTCTCCGACGACCCCGCGCTACATCTCTCC 666
QY 936 AGCTGAGATGGCAACGTTTACGCTGAGCTCGTTGACGCGTACCAATCGTCTCCATCGA 995
DB 667 CGAGCAGCTGGCGGACCTTACAGGGCTTCGTCAAAAGATTACCCAGTGTGTCCATCGA 726
QY 996 GGACCCACTGAGGAAGATGATCGGAGGTTTACCAACCTTACCGGAACCATTCGGCGGA 1055
DB 727 GGACCCCTTTTACGAGGACGACTGGGAGGCTGACCAATTTTACAGCCGACGACGAGC-- 784
QY 1056 CAAGGTTTCAGATCGTTGCGGACGACTTCTTCTGTCACCAACCTTACGAGCGCTGAAGGAGG 1115
DB 785 -----ATTGAGTGTGGGCGACGACCTTACGCTTACCAACCCCAACGCAATCGCAAGGG 840
QY 1116 CATCGTAAAGAGGCTGCAACTCCATCTGTTTAAAGTGAACAGATCGGTACCTCTCAC 1175
DB 841 TGTGGCCAGAGTCTTCAACTGCTGCTCTCAAGTCAACAGATCGGCTCGGTCTAC 900
QY 1176 CGAGACTTTCGACGCTGTGACATGCTGCTACCGCGAGGCTTACACTTCCATGATGTCCCA 1235
DB 901 CGAGTCTCTGAGGATGCAAGATGGCCGAGCAACGCGCTGGGCGGTGATGTTGTCAGCCA 960
QY 1236 CCGTTTCGGTGTGAGACGAGGACCAACATTTGCTGAGCTTGCAGTTTGCATCTCACTGTGG 1295
DB 961 TCGCTCCGGGAGACGAGGACACTTTCATCGCTGACCTTGTGCTGCTGTGTCACCGG 1020
QY 1296 CCAGATCAAGACTGGTGTCTCCAGCAGCTTCCGACCGGTGTGCGAAAGTACAAACGAGTTCT 1355

```

```

DB 1021 ACAGATCAATACGGGCGCTCCGTGCGCTCCGAGCGCTTGCCCAAGTACAAACGAGTCT 1080
QY 1356 CCGCATCGACGAGCTGTTTGGCGACGCGCGGCTTACGCGAGTTCGACGCGCATTTCCACG 1415
DB 1081 CCGATCGAGAGAGCTGGGCGACGACGCGGCTTCCCGCGCAAGAACTTCAGGACACCC 1140
QY 1416 CTTTCAGGG 1424
DB 1141 ATCTGAGCG 1149

RESULT 15
CR658829
LOCUS Tetraodon nigroviridis full-length cDNA.
DEFINITION Tetraodon nigroviridis full-length cDNA.
ACCESSION CR658829
VERSION CR658829.1 GI:51155274
KEYWORDS HTC; cDNA; full-length; Tetraodon nigroviridis.
SOURCE Tetraodon nigroviridis
ORGANISM Tetraodon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodontidae; Tetraodon.
REFERENCE 1 (bases 1 to 1551)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (10-AUG-2004) Genoscope - Centre National de Sequenage -
: 2 rue Gaston Cremlieux, CP 5706 - 91057 EVRY cedex - FRANCE
COMMENT The sequences are based on single pass reads.
More information available at
http://www.genoscope.cns.fr/tetraodon.
FEATURES
source
1..1551
/organism="Tetraodon nigroviridis"
/mol_type="mRNA"
/db_xref="taxon:99883"
/tissue_type="Muscle"

ORIGIN
Query Match 20.8%; Score 328; DB 3; Length 1551;
Best Local Similarity 57.9%; Pred. No. 2.8e-77;
Matches 740; Conservative 0; Mismatches 490; Indels 48; Gaps 7;

QY 160 ATATGACGATATTCGCTCGGAAATTCGATCCCGGTAACCCAAACGTCGAGGCA 219
DB 87 ATCACTAAGATCCAGGCTCGGAGATCTCGACTCCAGAGGAAACCCCAACGTTGGAGTG 146
QY 220 GAGGTTTCTCTGATGACGTTCCCGGTTGTCAGGTTGTCATCGGGGATCCACC 279
DB 147 GACTGTGGACGGCCAAAGG---CCTGTTCCAGACGACCGTCCCGGGGCTCTACT 203
QY 280 GCGCTCCACGAGGCTCATGAGCTGCGTGACGTTGGGCA---TGCCTACCTGGGCAAGGC 336
DB 204 GGGTCCATGAGGCGCTGGAGCTTCGTATGGAGACAGAGTTCGTACCTTGGCAAGGA 263
QY 337 GTTTTGAAGGAGTTGAAACGTCGCAAGAAATCGGCGACGAGTTCGCTG-----GC 390
DB 264 ACCCTGAAGGCGTGGACACGCTCAACAGGACATCGCCCCCAAGCTGATTGAGAAGAA 323
QY 391 CTAGAGGCTGACCATCAGCGCTTCATCGACGAAGCATGATCAAGTTGATGGCACCGCC 450
DB 324 TTGAGGCTTGTGGAGCAGGAGAGATTGACAGTTTATGCTGAGGCTGAGTGAACCGAG 383
QY 451 AACAAGTCCGCTGGGTGCAACGCAATCTCTTGGTGTTCATGCTGGTGTTCGAAAGGCT 510
DB 384 AACAAATCTCAGTTTGGGCGCAACGCAATCTTGGGCTGTCTCTGCGGCTCTGCAAGGCC 443
QY 511 GCTGCTGATTCGCGAGGCTTCCCACTGTTTCCGCTACATCGGTGAGCAACGACACG-- 568
DB 444 GGGGCGGGGAGAAAGGCTTCTCTCTACCGCCACATCGCCGATCTTCGCGCGGACAAA 503

```

Qy	569	-----TTCTTCAGTTTCCAAATGATGAAGAAATCATCAACGGTGGCGCTCAGCTGACTCC	621
Db	504	GACGTGATCCTCCCGGTTCCTGCTTTTCAAGGTATCAATCGTCGGAGGACCAACGCTGGGAAC	563
Qy	622	GGTGTTCAGCTTCAGGAATTCATGATCGCTCCAAATCGGTGACAGAGACCTTCTCTGAGGCT	681
Db	564	AAGCTGCCATATGCAAGGAGTTTCAATGTTCCCGCGTTGGAGCGGCCAACTTCCACGAGGCC	623
Qy	682	CTCCGAACCGCGCGGAGGTCTACCAACGCACTGAAGTCCGTTCATCAAGGAAAAAGGGC	738
Db	624	ATGAGGATCGCGCTGAGGTTTACCAAACTGAAGAGCGTGATCAAGGCAAAAGTACGGA	683
Qy	739	---CTGTCCACCGGACTTGGCGATGAGGGCGGCTTCGCTCTTCCTCGTCGCGCTCCACCCGT	795
Db	684	AAAGACCCCAACACGTGGGAGACGAAGGAGGCTTCGCCCCCAACATCTCTGSAGAAACAAC	743
Qy	796	GAGGCTCTTGACCTTATCGTTAAGGCAATCGAGAAGGCTGGCTTCACCCACGACGAAGGAC	855
Db	744	GAAGCTTGAGCTGCTGAAGTCAGCCATCGAGNAGCCGGCTACCCCGACAAGATCATC	803
Qy	856	ATCGCTCTTGCTCTGGAAGTTGCTTCCTCTGAGTTCTTTCAAGGACGGCACTTACCACCTTC	915
Db	804	ATCGGATGACGCTGGCGGCTCCGAGTTCTTATCGACGGGAAGTACGACCTGGACTTC	863
Qy	916	GA-----AGGTGGCCAGCACTCCGACGTCAGATGGCAACGTTTACGCT	960
Db	864	AAGTCCCCCGATGACCCATCCAGGCACATCAGCGGGAGAACTGGGAGACCTGTACCGC	923
Qy	961	GAGCTCGTTGACCGGTACCCAAATCGTCTCCATCGAGGACCCACTGCAAGGAAGATGACTGG	1020
Db	924	AGCTTCATCAGAACTACCTCTGTTTCAGTCCATCGAAGACCCCTTGACACGAGATGACTGG	983
Qy	1021	GAGGGTTACACCACTCACCGAACCATCGGGACAAAGTTTCAGATCGTTGGCGAGCAC	1080
Db	984	GAGAACTGGGCGAAGTTTACCGCTCCACAGACATC-----CAGATCGTTGGAGATGAC	1037
Qy	1081	TTCTTCTGTACCAACCCCTGAGCGCTCAAGGAGGGCATTCGCTAAGAGGGCTGCCAACTCC	1140
Db	1038	CTGACGGTGCACCAACCCCAAGAGATCCAGACGGCGTGCACAAGAAGGCCCTGTAACTGT	1097
Qy	1141	ATCCTGTTTAAAGTGAACAGATCGGTACCTTCACCGAGACCTTCGACGCTGTGCACATG	1200
Db	1098	CTGCTGTCAAGGTCAACCAAGATCGGCTCAGTGAACCGAGTCCATCAAGCGGTGTAAAGCTG	1157
Qy	1201	GCTCACCGCGCAGGCTACACCTTCATATGTCCACACCGTTCCGCTGAGACCGGAGGACACC	1260
Db	1158	GCCCAGAGCAGTGGTTGGGGTGTGATGGTCAGCCATCGCTCTGGAGAACTGAGGACACC	1217
Qy	1261	ACATTCCTGACCTTCGGAGTTGCACTCAACTGTGTGGCCAGATCAAGATCGTGTCTCAAGCA	1320
Db	1218	TTTCATCTCCGACCTTGGTGTGGACTGTGACCGGACAGATCAAGACGGGCGCTCCCTGC	1277
Qy	1321	CGTTCCACCGCTGTGCAAAAGTCAACACAGCTTCTCCGCATTCGAGCAGCTGCTTGGCGCAC	1380
Db	1278	AGGTGAGGCGGTTTGGCCAAAGTACACACGCTGATGAAATCGAGGAAGAGCTTGGGGAC	1337
Qy	1381	GCGGGCGTCTACGAGGT	1398
Db	1338	AAGGCCAAGTTCCCGGT	1355

Search completed: September 28, 2005, 23:24:12
Job time : 4763.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 28, 2005, 16:08:34 ; Search time 275.5 Seconds
(without alignments)
9372.213 Million cell updates/sec

Title: US-10-728-947-3
Perfect score: 1578
Sequence: 1 ggctgggatatggtagtt.....ctcaagcagggaacgtgctt 1578

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PTUS_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1578	100.0	1578	4	US-09-860-768-3
2	1576.4	99.9	1578	4	US-09-860-768-1
3	632.6	40.1	4403765	3	US-09-103-840A-2
4	632.6	40.1	4411529	3	US-09-103-840A-1
5	550.4	34.9	1299	4	US-09-902-540-4780
6	550.4	34.9	27903	4	US-09-902-540-1235
7	532	33.7	1335	4	US-09-252-991A-4449
8	531.4	33.7	1347	4	US-09-252-991A-4158
9	479.6	30.4	1380	4	US-09-489-039A-2890
10	449.2	28.5	2363	4	US-09-818-780-22
11	440.8	27.9	3545	2	US-08-781-802-9
12	431.6	27.4	1305	4	US-09-583-1102-1156
13	430.4	27.3	11384	3	US-08-961-527-45
14	428	27.1	1196	4	US-09-107-532A-156
15	418.6	26.5	1233	4	US-09-489-039A-2858
16	414	26.2	1511	3	US-08-847-065-18
17	414	26.2	1511	4	US-09-829-382-18
18	390.2	24.7	3513	3	US-09-058-260-1
19	390.2	24.7	4315	3	US-08-781-802-1
20	390.2	24.7	4315	3	US-08-694-078-1
21	389.6	24.7	1830121	4	US-09-557-884-1
22	389.6	24.7	1830121	4	US-09-643-990A-1
23	383.8	24.3	1371	4	US-09-543-681A-1918
24	381.6	24.2	1224	4	US-09-107-433-2383
25	374.8	23.8	1335	4	US-09-328-352-3143
26	367.2	23.3	1338	4	US-09-540-236-434
27	367.2	23.3	92407	4	US-09-596-002-36

28	359.6	22.8	1460	4	US-09-830-217-21	Sequence 21, Appl
29	359.6	22.8	1460	4	US-10-278-946-21	Sequence 21, Appl
30	351.2	22.3	1305	4	US-09-710-279-2221	Sequence 2221, Ap
31	351.2	22.3	1323	3	US-09-134-001C-2804	Sequence 2804, Ap
32	351.2	22.3	3830	4	US-09-710-279-3773	Sequence 3773, Ap
33	344.6	21.8	1539	4	US-08-956-171B-207	Sequence 207, App
34	344.6	21.8	1539	4	US-08-781-986A-207	Sequence 207, App
35	340.6	21.6	1230230	4	US-09-438-185A-1	Sequence 1, Appli
36	330.2	20.9	1230025	4	US-09-198-452A-1	Sequence 1, Appli
37	329	20.8	1664976	4	US-08-916-421B-1	Sequence 1, Appli
38	329	20.8	1664976	4	US-09-692-570-1	Sequence 1, Appli
39	308	19.5	1314	4	US-09-614-221A-509	Sequence 509, App
40	304.8	19.3	1314	4	US-09-614-221A-466	Sequence 466, App
41	304.8	19.3	1884	4	US-09-566-921-38	Sequence 38, Appl
42	301	19.1	934	3	US-08-961-083-183	Sequence 183, App
43	301	19.1	934	4	US-09-536-784-183	Sequence 183, App
44	299.6	19.0	2422	4	US-09-949-016-282	Sequence 282, App
45	299.6	19.0	2422	4	US-09-949-016-2488	Sequence 2488, App

ALIGNMENTS

RESULT 1
US-09-860-768-3
; Sequence 3, Application US/09860768
; Patent No. 6713289
; GENERAL INFORMATION:
; APPLICANT: Mockel, Bettina
; APPLICANT: Pfeifferle, Walter
; APPLICANT: Hermann, Thomas
; APPLICANT: Pohler, Alfred
; APPLICANT: Kalinowski, Jorn
; APPLICANT: Bathe, Brigitte
; TITLE OF INVENTION: New Nucleotide Sequences that Code for the Eno Gene
; FILE REFERENCE: 21123/278404
; CURRENT APPLICATION NUMBER: US/09/860,768
; CURRENT FILING DATE: 2001-05-21
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 3
; LENGTH: 1578
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
US-09-860-768-3

Query Match	100.0%	Score	1578	DB	4	Length	1578
Best Local Similarity	100.0%	Pred. No.	0				
Matches	1578	Conservative	0	Mismatches	0	Indels	0
Gaps	0						
Qy	1	GGCTGGGGATATGGGTAGTTTTCGCCACTTAATTTCAACTGATTCCTCATCGAAACAAGA	60				
Db	1	GGCTGGGGATATGGGTAGTTTTCGCCACTTAATTTCAACTGATTCCTCATCGAAACAAGA	60				
Qy	61	TTCTGTGCAACAATTTGGGTAGTACGATTTGAAGACATTTTGATCAGTGAATTAATTCCTAG	120				
Db	61	TTCTGTGCAACAATTTGGGTAGTACGATTTGAAGACATTTTGATCAGTGAATTAATTCCTAG	120				
Qy	121	TTAGTCTCCCAAGTTGGCATAGGAGGCCACAGTGGTGAATCATGACATTCGTCGTCGC	180				
Db	121	TTAGTCTCCCAAGTTGGCATAGGAGGCCACAGTGGTGAATCATGACATTCGTCGTCGC	180				
Qy	181	GAAATTCGACTCCCGCGGTACCAACCGTCGAGCAGAGTTTTCCTGGATGACGGT	240				
Db	181	GAAATTCGACTCCCGCGGTACCAACCGTCGAGCAGAGTTTTCCTGGATGACGGT	240				
Qy	241	TCCACCGGTGTCGAGGTGTTCCATCCGGCGCATCCACCGCGCTCCACGAGGTCATGAG	300				
Db	241	TCCACCGGTGTCGAGGTGTTCCATCCGGCGCATCCACCGCGCTCCACGAGGTCATGAG	300				
Qy	301	CTCGGTGACGGTGGCGATCGCTACCTGGGCAAGGGCGTTTGAAGGAGTGAACAGTC	360				
Db	301	CTCGGTGACGGTGGCGATCGCTACCTGGGCAAGGGCGTTTGAAGGAGTGAACAGTC	360				

Db	481	CTTGGTGTTCATGGCTGTGGAAAGGCTGCTGCTGATTTCCGCAGGCGCTCCCACTGTTTC	540
Qy	541	CGCTACATCGGTGGACCAAAACGGACACAGTTCCTCCAGTTCCAAATGATGAACATCATCAAC	600
Db	541	CGCTACATCGGTGGACCAAAACGGACACAGTTCCTCCAGTTCCAATGATGAACATCATCAAC	600
Qy	601	GGTGGCGCTCAGCGTGACTCGGTGGTGAAGTTCAGGAAATTCATGATCGCTCCATCGGT	660
Db	601	GGTGGCGCTCAGCGTGACTCGGTGGTGAAGTTCAGGAAATTCATGATCGCTCCATCGGT	660
Qy	661	GCAGAGACCTTCTCTGAGGCTCTCCGCAAACGGGCGGAGGTCTACACGACACTGAAGTCC	720
Db	661	GCAGAGACCTTCTCTGAGGCTCTCCGCAAACGGGCGGAGGTCTACACGACACTGAAGTCC	720
Qy	721	GTCAATCAAGGAAAGGCGCTGTCCACCGGACTTGGCGGATGAGGCGGCTTCGCTCCTTCC	780
Db	721	GTCAATCAAGGAAAGGCGCTGTCCACCGGACTTGGCGGATGAGGCGGCTTCGCTCCTTCC	780
Qy	781	GTGGCTCCACCGTGAGGCTCTTTGACCTTATCGTTAAGGCAATCGAAGAGGTGGCTTC	840
Db	781	GTGGCTCCACCGTGAGGCTCTTTGACCTTATCGTTGAGGCAATCGAAGAGGTGGCTTC	840
Qy	841	ACCCAGGCAAGGACATCGCTCTTTGCTCTGACGCTGCTTCTCTGAGTTCCTTCAAGGAC	900
Db	841	ACCCAGGCAAGGACATCGCTCTTTGCTCTGACGCTGCTTCTCTGAGTTCCTTCAAGGAC	900
Qy	901	GGCACTTACCACTTTCGAAGGTGGCCAGCACTCCGACGCTGAGATGGCAAAAGTTCAGCT	960
Db	901	GGCACTTACCACTTTCGAAGGTGGCCAGCACTCCGACGCTGAGATGGCAAAAGTTCAGCT	960
Qy	961	GAGCTCGTTGACCGGTACCCAAATCGTCTCCATCGAGGACCCACTGCGAGGAAGATGATCG	1020
Db	961	GAGCTCGTTGACCGGTACCCAAATCGTCTCCATCGAGGACCCACTGCGAGGAAGATGATCG	1020
Qy	1021	GAGGTTTACACCACTTACCGCAACCATCGGCGACAAGTTTCAGATCGTTGGCGAGCAC	1080
Db	1021	GAGGTTTACACCACTTACCGCAACCATCGGCGACAAGTTTCAGATCGTTGGCGAGCAC	1080
Qy	1081	TTCTTTCGTACCAACCTTGAGCGGCTCGAAGGAGGCAATCGCTAAGAAGGCTGCCAATCC	1140
Db	1081	TTCTTTCGTACCAACCTTGAGCGGCTCGAAGGAGGCAATCGCTAAGAAGGCTGCCAATCC	1140
Qy	1141	ATCTTGGTTAAGGTGAACAGATCGGTACCCCTCAACGAGACCTTTCGACGCTGTGCACATG	1200
Db	1141	ATCTTGGTTAAGGTGAACAGATCGGTACCCCTCAACGAGACCTTTCGACGCTGTGCACATG	1200
Qy	1201	GCTCACCGCGAGGCTACCTCCATGATGTCCACCGTTCCGCTGAGACCGAGGACACC	1260
Db	1201	GCTCACCGCGAGGCTACCTCCATGATGTCCACCGTTCCGCTGAGACCGAGGACACC	1260
Qy	1261	ACCAATGCTGACCTCGCAGTTTGCACTCAACTGTGGCCAGATCAAGACTGGTGTCTCCAGCA	1320
Db	1261	ACCAATGCTGACCTCGCAGTTTGCACTCAACTGTGGCCAGATCAAGACTGGTGTCTCCAGCA	1320
Qy	1321	CGTTCCGACCGGTGTCGAAAGTCAACACAGCTTCTCCGCACTCGAGCAGCTGCTTGGCGAC	1380
Db	1321	CGTTCCGACCGGTGTCGAAAGTCAACACAGCTTCTCCGCACTCGAGCAGCTGCTTGGCGAC	1380
Qy	1381	GC CGCGGTCTACG CAGGTTCGAGCGCAATTCCACGCTTTTCAGGCGCTAAATAAAGCGCTT	1440
Db	1381	GC CGCGGTCTACG CAGGTTCGAGCGCAATTCCACGCTTTTCAGGCGCTAAATAAAGCGCTT	1440
Qy	1441	TTCCAGCGCCGGTAACTCAAGGTTCGCGGCGCTGCTGCTTACTACTGTTACTGCTGT	1500
Db	1441	TTCCAGCGCCGGTAACTCAAGGTTCGCGGCGCTGCTGCTTACTACTGTTACTGCTGT	1500
Qy	1501	GACTATGATCGAGGATTTATGGCAAGCAGAGAAAACTATAAAGGCGCTTGTTCCTGCTCT	1560
Db	1501	GACTATGATCGAGGATTTATGGCAAGCAGAGAAAACTATAAAGGCGCTTGTTCCTGCTCT	1560
Qy	1561	CAAGCAGGGAACGTGCTT 1578	

Db 1561 CAAGCAGGGAACTGTCTT 1578

RESULT 3

US-09-103-840A-2

; Sequence 2, Application US/09103840A

; Patent No. 6294328

; GENERAL INFORMATION:

; APPLICANT: FLEISCHMAN, Robert D.

; APPLICANT: WHITE, Owen R.

; APPLICANT: FRASER, Claire M.

; APPLICANT: VENTER, John C.

; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM

; TITLE OF INVENTION: TUBERCULOSIS

; FILE REFERENCE: 24366-20007.00

; CURRENT APPLICATION NUMBER: US/09/103,840A

; CURRENT FILING DATE: 1998-06-24

; NUMBER OF SEQ ID NOS: 2

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 2

; LENGTH: 4403765

; TYPE: DNA

; ORGANISM: Mycobacterium tuberculosis

; FEATURE:

; OTHER INFORMATION: CDC 1551

; OTHER INFORMATION: "n" bases at various positions throughout the sequence

; OTHER INFORMATION: represent a, t, c or g

US-09-103-840A-2

Query Match 40.1%; Score 632.6; DB 3; Length 4403765;

Best Local Similarity 68.4%; Pred. No. 8.4e-166;

Matches 891; Conservative 0; Mismatches 409; Indels 3; Gaps 1;

Qy 140 AGGAGGCCACAGTGGCTGAAATCATGCACGTATTCGCTCGCGAAATTTCTCGACTCCCGCG 199

Db 1144576 AGNGAACAGTGGCGATTATCGAGCAGTTAGGCCCCGAGAGATCTCGATTCCTCGCG 1144635

Qy 200 GTAAACCAACCGTCGAGCGAGAGTTTCTGTGATGACGGTTCCACGGTGTTCGAGGTG 259

Db 1144636 GCAACCCGACGGTGGAGTCCAGTGGCGCTTATCGACGGGACATTCGCCCGGCCGCGG 1144695

Qy 260 TTCCATCCGGCGCATCCACGGCGTCCACAGGCTCATGAGCTCGTGAGTGGGCGATC 319

Db 1144696 TGCCTGCGGCGCTCGACCGGGAGCAGAGGCCCTCGAGTTGCGCGACGGCGGCGATC 1144755

Qy 320 GCTACCTGGCAAGGGCGTTTCAAGCGAGTTGAAACCGTCAACGAAGAAATCGGGACG 379

Db 1144756 GCTACGGCGCAAGGGGTCAAAAGCCGTGCAGGCTGTTCTTGATGAGATCGGCCGG 1144815

Qy 380 AGTTCGCTGCGCTAGAGGTGACGATCAGCGCTCATCGACGAAGCAATGATCAAGCTTG 439

Db 1144816 CCGTCATCGGACTCAACCGCGACGACAGCGATTTGGTCGACACGAGGCGTGGTGACCTAG 1144875

Qy 440 ATGGCAGCCCAACAGTCCGGCTGGTGCAACCGAATCTTGGTGTTCATCGCTG 499

Db 1144876 ACGGCACCCCGCAAGTCCCGCTGGGGGCAACCGGATCTTTGGGTGTCCTCGCTG 1144935

Qy 500 TTGCAAGGGTGTGCTGATTTCGCGAGGCGCTCCACCTGTTCCGCTACATCGGTGGACCAA 559

Db 1144936 TTGCAAGGGCGGGCGGATTCGCGAGAGTCCGTTGTTCCGTTATGTCGGGGGGCCAA 1144995

Qy 560 ACGCACACGTTCTTCCAGTTTCCAATGATGAACATCATCAACGGTGGCGCTCACGCTGACT 619

Db 1144996 ACGGCACATTCGCGGTACCGGATGATGAACATCCTCAACGGCGCGCACACGCCGATA 1145055

Qy 620 CCGGTGTTGACGTTTCAGGAATTCATGATCGCTCAATCGGTGCAGAGACCTTCTCTGAGG 679

Db 1145056 CCGGTGTCGACATTCAGAGTTTCATGTTGGTGGCGCAATTTGGCGGCGCCAGCTTCGTGAGG 1145115

Qy 680 CTCTCCGCAACGGCGGAGGTGTACACGACCTGGAAGTCCGTCATCAAGGAAAGAGGGCC 739

Db 1145116 CGTTGCGCTGGGCGCTGAGGTGTACACGCGTCAAGTCCGTCCTGAAAGAGAGGGGC 1145175

```
QY 740 TGTCCACGGACTTGGCGATGAGGCGGCTTCGTCTTCGTGGCTCCACCGGTGAGG 799
Db 1145176 TGTCCACGGCTCGGCGACGAAGCGGCTTCGCCCGGATGTGCCGCGACACCGCGG 1145235
QY 800 CTCTTGACCTTATCGTTAAGGCAATCGAGAAGGCTGGCTTCACCCGAGGCAAGACATCG 859
Db 1145236 CGTTGGACCTGATCAGCGCGGCCATCGAGTCGGCGGGCTTCGACACCGCGCGACGTGG 1145295
QY 860 CTCTTGCTCTGGAGTGTCTTCCTCTGAGTCTTCAAGGAGGAC---CTACACTTCG 916
Db 1145296 CGTGGCCCTTGGACCGCGCGGCCACCGAGTTCTTCACCGGACCGGCTACGCTCTCG 1145355
QY 917 AAGTGGCCAGCACTCCGCGAGCTGAGATGGCAACGTTTACGCTGAGCTGTTGACGCGT 976
Db 1145356 AGGCAACACCGTACCGGACGACGATGACCGAGTTCTACCGGGCTGCTCGCGCCT 1145415
QY 977 ACCCAATGCTTCATCGAGGACCACTGCGAGGAAGATGACTGGGAGGTTTACACCAACC 1036
Db 1145416 ACCCGCTGGTGCATCGAAGACCCACTGTCCGAAGACGATTGGGACGGCTGGGCGCGC 1145475
QY 1037 TCACCGCAACCATCGGCGACAGGTTGAGATCGTTGGGAGGACTTCTTCGTACCAACC 1096
Db 1145476 TGACGGCTCGATCGGTGACCGGGTGCAATCGTGGCGGACGACATCTTTGTACCAATC 1145535
QY 1097 CTGAGCGCTGAAGGAGGATCGCTAAGAAGGCTGCCAATCCATCTCGTTAAGGTGA 1156
Db 1145536 CCGAGCGCTCGAGGAGGATCGAACGGGCGTGGCAATGCTGCTCGTCAAGGTGA 1145595
QY 1157 ACAGATCGGTACCTCACCGAGACCTTCGACGCTGTCGACATGGCTCACCGCGAGGCT 1216
Db 1145596 ACCAGATCGGAGCTTGACCGAGACACTCGACGCGGTCAAGCTGGCTCACCGCGGAT 1145655
QY 1217 ACACCTCATGATGTCACCGCTTCGCTGAGACCGGAGGACACCAATTCGTGACCTCG 1276
Db 1145656 ACCGACGATGATCAGTCACCGAGTGGCGAGACGAGGACACATGATCGCGACCTCG 1145715
QY 1277 CAGTTGACCTCAAGTGTGGCCAGATCAAGACTGCTGTCAGACAGCTTCGACCGTTCG 1336
Db 1145716 CGGTGGCCATCGGACGGGCGAGTCAAGACGGGCGGCTGCTCGAGTGAGGCGTGC 1145775
QY 1337 CAAAGTACACAGCTTCTCCGATCGAGCAGCTGCTTGGCGAGCGCGGCTTACGCGAG 1396
Db 1145776 CAAATAACAAACGACTGCTCGGATCGAAGAGGCGCTTGGCGAGCGCGGCGCTACGCG 1145835
QY 1397 GTCGACGCGATTCACACGCTTCAGGCGCTAAATAAAGCGCT 1439
Db 1145836 CGACCTGGCATTTCTCGGTTCCGCTGCGAGAGCAATAGGT 1145878
```

RESULT 4

```
US-09-103-840A-1
; Sequence 1: Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 441529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1
```

Query Match

40.1%; Score 632.6; DB 3; Length 4411529;

```
Best Local Similarity 68.4%; Pred. No. 8.4e-166;
Matches 891; Conservative 0; Mismatches 409; Indels 3; Gaps 1;

QY 140 AGGAGGCCACAGTGGGTGAAATCATGCACGATATTCGCTCGCGAAATTTCTCGACTCCCGCG 199
Db 1144551 AGGAGAACCCAGTGCAGATTATCGAGCAGGTTAGGGCCCGAGAGATCCTCGATTCCGCG 1144610
QY 200 GTAACCAACCGTCGAGGCGAGAGTTTTCTTGATGACGGTTCCCAAGGTTGTCGAGGTG 259
Db 1144611 GCAACCCGACGCTGGAGGTGAGGTGGCGCTTATCGACGGGACATTCGCCCGGCGCGCG 1144670
QY 260 TTCATCTCGGCGCATCCACGGCGTCCACGAGGCTCATGAGCTCGCTGAGCGTGGCGATC 319
Db 1144671 TGCCTGCGGCGCTCCACCGGGAGCACGAGGCGCTCGAGTTGCGGACGCGGCGCATC 1144730
QY 320 GCTACCTGGGCAAGGGCGTTTTTGAAGGCAAGTTTGAACACGTCACAGAAATCGGCGACG 379
Db 1144731 GCTACGGCGCAAGGCGTGCAAAAAGCGTGCAGGCTGTTCTTGATGAGATCGGCCGG 1144790
QY 380 AGCTCGCTGGCCTTAGAGGCTGACGATCAGCGCTCATTCGACGAAGCAATGATCAAGCTTG 439
Db 1144791 CCCTCATCGGACTCAACGGCGACGACGAGATTGGTTCGACCAAGGCGCTGGTGACCTAG 1144850
QY 440 ATGGCACCGCAACAAAGTCCCGCTCGGTGCAACGCAATCTCTTGGTGTTCATCGCTG 499
Db 1144851 ACGGCAACCCCGCAAGTCCCGCTGGGCGCAACGCGATCTTGGGTGTCTCGCTCGCTG 1144910
QY 500 TTGCAAGGCTGCTGCTGATTCGCGAGGCTCCCACTGTTTCGCTACATCGGTGGACCAA 559
Db 1144911 TTGCCAAGCGCGCGCGGATTCGCGGAGCTGCGGTGTTCCGTTATGTGCGGGGCCAA 1144970
QY 560 ACGCACACGTTCTTCAGTTTCCAATGATGAACATCATCAACGTTGGGCTCAGCTGACT 619
Db 1144971 ACGCGACATTTTCGCGGATCCGCGATGAACATCTCTCAACGGCGCGCACACGCCGATA 1145030
QY 620 CGGCTGTTCAAGTTTCAAGAAATTCATGATCGTCCAATCGGTGCGAGAGACTTCTCTGAGG 679
Db 1145031 CGGCTGTCACATTCAGAGTTTCATGTTGGCGGCAATTTGGCGGCCAGCTTCGTGAGG 1145090
QY 680 CTCCTCGCAACGGCGCGGAGGTCTACCAAGCTGAAAGTCCGTCATCAAGAGAAAGGGCC 739
Db 1145091 CGTTGCGCTGGGCGCTGAGGTGTACACGCGCTCAAGTCGCTCTGAAAGAGGAGGGGC 1145150
QY 740 TGTCCACCGGACTTGGCGATGAGGCGGCTTCGCTCTTCGCTCGGCTCCACCGTGAAG 799
Db 1145151 TGTCCACCGGCTTGGGCGACGAAGGCGCTTCGCCCGGATGTGGCGGCGCACCGCGCG 1145210
QY 800 CTCCTGACCTTATCGTTAAGGCAATCGAGAAGGCTGGCTTCAACCCGAGGCAAGGACATCG 859
Db 1145211 CGTTGACCTGATCAGCGCGGCCATCGAGTGGCGGCTTTCGACCGGCGCGACGTGG 1145270
QY 860 CTCCTGCTCTGAGCGTTGTTCTCTGAGTTCTTCAAGAGCGGAC---CTACCACTTCG 916
Db 1145271 CGTGGCCCTTGGACCGCGGCGCACCGAGTTCTTCACCGACGCGACCGGCTAGCTCTTCG 1145330
QY 917 AAGTGGCGGACACTCCGCGAGCTGAGATGGCAACGTTTACGCTGAGCTCGTTGACGCGT 976
Db 1145331 AGGCAACACCGTACCGGACACGATGACCGAGTTCTACCGGGGCTGCTCGGCGCT 1145390
QY 977 ACCCAATCGCTCCATCGAGGACCCACTGCGAGGAAGATGACTGGGAGGTTTACACCAACC 1036
Db 1145391 ACCCGTGTGTGTCATCGAAGACCCACTGTCCGAGAGCAGATTGGGACGCTGGGCGCGCG 1145450
QY 1037 TCACCGCAACCATCGGCGACAGGTTTCAAGTTCAGTGTGGCGACGACTTCTTCGTCAACCAACC 1096
Db 1145451 TGACGGCTTCGATCGGTGACCGGGTGCAATCGTGGCGGACGACATCTTTGTCAACCAATC 1145510
QY 1097 CTGAGCGCTGAAGGAGGCGCATCGCTAAGAAGCGTCCCAACTCCATCTCTGTTTAAAGTGA 1156
Db 1145511 CCGAGCGGCTCGAGGAGGCGCATCGAACGGGCGTGGCAATGCGTTGCTTCTGTTCAAGGTGA 1145570
QY 1157 ACCAGATCGGTACCTTCACCGAGACTTCGACGCTTCGACATGGCTCACCGCGAGGCT 1216
```

Db 1145571 ACAGATCGGAGCTTTCACCGAGACATCGACGCGGTCAACGCTGGCTCAACACGCGCGAT 1145630
Qy 1217 ACACCTCCATGATGTCCTCCACCGTTCGGGTGAGACCGGAGGACACCAACATTCGTGACCTCG 1276
Db 1145631 ACCGACGATGATCAGTCAACGAGTGGCGAGAGGAGACACCATGATCGCCGACCTCG 1145690
Qy 1277 CAGTTGCACTCAACTGTGGCCAGATCAAGACTGTGTGTCAGACAGCTTCGACCGTGTG 1336
Db 1145691 CGGTGGCCATCGGACGCGGCGAGATCAAGACGCGCGCGCTGTCTCGCAGTGAGCGCTCG 1145750
Qy 1337 CAAAGTACACCACTTCTCCGATCGAGCAGCTGTGGGAGCGCGGTCTACGAG 1396
Db 1145751 CAAATAACACCACTGCTGCGGATCGAAGAGGCGCTTGGCGAGCGCGCGCTACGCG 1145810
Qy 1397 GTGCGACGCAATCCCAACGCTTTTCAGGCTTAAATAAAGCGCT 1439
Db 1145811 GCGACCTGGCAATTCCTCGGTTCGCGTGCGAGAGCAATAGGT 1145853

RESULT 5

US-09-902-540-4780
; Sequence 4780, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 4780
; LENGTH: 1299
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-4780

Query Match 34.9%; Score 550.4; DB 4; Length 1299;
Best Local Similarity 65.6%; Pred. No. 1.6e-144;
Matches 841; Conservative 0; Mismatches 426; Indels 15; Gaps 2;

Qy 152 TGCTGAATCATGCAGTATTCGCTCGCAATTCCTCGACTCCGCGGTAAACCAACCG 211
Db 2 TGACCGAGATTCTCAGATTGGCGCGCGAAGTGCTCGACTCCCGGTGGCAACCCGACCG 61
Qy 212 TCAGAGGAGAGGTTTCTCGATGACGCTTCCACGCGTTCGACAGGTGTTCCATCCGCGG 271
Db 62 TGGAGGCGAGGTTCCAGCTTGGCGGGGCGCTCGTGCGCGCGCGGTGCGGTCCGCTG 121
Qy 272 CATCCACCGCGTCCACGAGCTCATGAGTGTGCTGACGTTGCGG---GATCGTACTGG 328
Db 122 CGTCCACCGGTGAGCATGAGCCATCGAGTGTGCGGACGCGGCAACGATCGCTACCTGG 181
Qy 329 GCAAGGCGCTTTGAAGCGAGTTGAAACGTCACGAGAAATCGGCGAGCTGCTCG 388
Db 182 GCAAGGCGGTGCAAGGCGGTGAAGAACGTGTGTCGAGCTGCTCGCGCGCGCTGTGG 241
Qy 389 GCCTAGAGGTGACGATCAGCGCTCATCGACGAGCAATGATCAAGCTTGATGGCACCG 448
Db 242 GCATGAGCGCGCGGATCAGTTCCGGGTGACACGAGCGATGCTGGAGCTGGACGACCG 301
Qy 449 CCAACAGTCCCGCTCGGTGCAACCAATCTTGGTGTTCATCGCTGTGTCAAAGG 508
Db 302 CCACCAAGGCAAGCTGGCGGCCAACCCATCTTCGCGGTGTCTCATGGCGCGCTCGCG 361
Qy 509 CTGCTGCTGATTCGCGAGGCTCCCATGTTCCGCTACATCGGTGGACCAACGACAG 568
Db 362 CCGCGCGGACGCGACGCGGTGCGCTGTACCGTACGTGGGCGGTGCGAGCGCGCA 421

Qy 569 TTCTTCAGTTCCAATGATGAACATCATCAACGCTGCGCTCAGCTGATCTCCGGTGTG 628
Db 422 CCCTGCGGTGCGCTGATGAACATCTCAACGCGCGCGCACCGGACACCGCGGTG 481
Qy 629 ACCTTCAGGATTCATGATCGCTCCAATCGGTGCGAGACCTTCTCTGAGGCTCTCCGCA 688
Db 482 ACCTGAGGATTCATGCTGCTGCGCGCGGTGCTCTCTCTGCGGAGGCGCTCGGCT 541
Qy 689 ACAGCGCGGAGGTCTACACGCACTGAAGTCCGTCATCAAGGAAAGGCGCTGTCCACCG 748
Db 542 GGGCGCGGAGGTTCACGCGCTGAAGAATCTCAAGGGCGGCAAGCTGGCCACCG 601
Qy 749 GACTTGGCGATGAGGGCGGTTCGCTCTTCGCTGCGCTCCACCGGTGAGGCTCTTGACC 808
Db 602 GCGTGGCGAGAGGGCGGTACGCGCGGACCTGCGCGGCAACGAAGAGCGCTGAAGC 661
Qy 809 TTATCGTTAAGCAATCGAAGAGGCTGCTTCAACCGGAGGAGGATCGCTCTTGCTC 868
Db 662 TCATCATTGGAGGCGCATCGACGCGGCTTCAAGGCGGTGAGCAGCTTCTCTGGGCC 721
Qy 869 TGGACGTTGCTTCTCTGAGTCTTCAAGGACGCGCACTTACCACTT-----CG 916
Db 722 TGGACGTTGGCGCGCAGCGAGTCTTTCGACAAAGGCGCAGAGTACAAAGTGAAGGCG 781
Qy 917 AAGTGGCCAGCACTCCGAGCTGAGATGGAACGTTTACGTGAGCTGCTGACGCT 976
Db 782 AGGCAAGGAGTACGACTCGACCGCGCTGCTCGAGTACTACCGGGGCTCTCCGAGCGCT 841
Qy 977 ACCCAATCGTCTCCATCGAGGACCCACTCGCAGGAGATGATGGAGGCTTACACCAAC 1036
Db 842 ACCCAATCATCTCCATCGAAGAGCGCATGGCGGAGGATGATGGAGGCTTGAAGAAC 901
Qy 1037 TCACCGCAACCATCGGCGCAAGGTTTCAGATCGTTGGCGAGCTTCTTCGTCAACCAAC 1096
Db 902 TCACCGAGCGCTGGGTTCGCGCATGCAAGTGGTGGGTGAGCACTCTTCGTCAACCAAC 961
Qy 1097 CTGAGCGCTTGAAGAGGCGCATCGCTAAGAGGCTGCCAATCCATCTGTTAAGGTGA 1156
Db 962 TGGAGCGCTGGCGCGCGCATCGAGACGCGCACTCCATCTTCGTGAAGGTGA 1021
Qy 1157 ACAGATCGGTACCTCACCAGACCTTCAGCGCTGCGACATGGCTCACCGCGAGCT 1216
Db 1022 ACCAGATTTGCACTCCATCGAGGAGAGCTTCGACGCGGTGCGCATGGCGCACCGCGCT 1081
Qy 1217 ACACCTCATGATGTCACCGCTTCGCGTGAGACCGGAGGACACCACTTCGTGACCTCG 1276
Db 1082 ACAGTCCGTGATGAGCCACCGCTCCGCGGAGACGAGGAGACCAACCATCGCCGACCTGG 1141
Qy 1277 CAGTTGCACTCAACTGTGGCCAGATCAAGACTGCTCCAGACAGTTTCGACCGTGTG 1336
Db 1142 CCGTGGCGCTGGACTTCGCGGCGAGATCAAGACGGGTTCCGCGTCCCGCTCCGACCGCTGG 1201
Qy 1337 CAAAGTACACAGCTTCTCCGATCGAGCAGCTGCTTGGGAGCGCGGCTCTACGAG 1396
Db 1202 CCAAGTACAAACAGTTCGTCGATCGAAGCGAGCTGGGGGCGCGCGCTACGCTG 1261
Qy 1397 GTCGAGCGCATTCCTCCACGCTT 1418
Db 1262 GCAAGTGGTGTTCGCGCGCTT 1283

RESULT 6

US-09-902-540-1235/c
; Sequence 1235, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540

; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 1235
; LENGTH: 27903
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-502-540-1235

Query Match 34.9%; Score 550.4; DB 4; Length 27903;
Best Local Similarity 65.6%; Pred. No. 8e-144;
Matches 841; Conservative 0; Mismatches 426; Indels 15; Gaps 2;
QY 152 TGCGTGAATCATGCAGATATTCGCTCGGAAATTCCTGACTCCCGGGTAAACCCACCG 211
Db 19948 TGACCCGAGATTCTTCAGATTCTGGCGCGCAAGTGTCTGACTCCCGTGGCAACCCGACCG 19889
QY 212 TCAGGCGAGAGGTTTCTCGATGACGCTTCCACGCTGTCGAGGTGTTCCATCCCGCG 271
Db 19888 TGAGGCGGAGGTTCCAGCTTGCAGGCGGCGCTCGTGCCCGCGCGCGGTGCCGTTCGCGTG 19829
QY 272 CATCCACCGCGCTCCAGAGGCTCATGAGCTGGCTGACGGTGGC--GATCGCTACCTGG 328
Db 19828 CGTCCACCGGTGAGCATGAGGCCATCGAGCTGCGGAGCGGCAAGCATCGCTACTGG 19769
QY 329 GCAAGGCGGTTTGAAGGCAAGTTGAAACGTCACGGAAGAAATCGGGCGACGAGTCCGCTG 388
Db 19768 GCAAGGCGGTGCAAGAGCGCGTGAAGAACGTGTGTCGACGTGCTCGCGCCCGCGCTGGTGG 19709
QY 389 GCCTAGAGGCTGACGATCAGCGCTCATCGACGAAGCAATGATCAAGCTTGATGGCACCG 448
Db 19708 GCATGACGCGGCGATCAGTTCCGGTGGACACAGCGGATGCTGGAGCTGGACGGCACCG 19649
QY 449 CCAACAGTCCCGCTGGGTGCAAAACGCAATCTTGTGTGTTCCATGGCTGTGTCGAAAGG 508
Db 19648 CCACCAAGGCAAGCTGGCGCGCAACGCGCATCTCGCGGTGTCCATGGCGCGCTCGCG 19589
QY 509 CTGCTGCTGATTCGGCAGGCTCCCACTGTTCCGCTACATCGGTGGACCAAAACGACAG 568
Db 19588 CCGCGCGGACGGCGACGGGCTCCGCTGTACCGCTACGTGGGCGGCGTGCAGCGCGCA 19529
QY 569 TTCTTCCAGTTCCAATGATGAACATCATCAACGCTGGCGCTCAGCTGACTCCGGTGTG 628
Db 19528 CCTGCGCGTCCGCTGATGAACATCTCAACGGCGCGCGACGCGGACACCGCGTGG 19469
QY 629 ACCTTCAGGAATTCATGATCGCTCCATCGGTGCAGAGACCTTCTCTGAGGCTCTCCGCA 688
Db 19468 ACCTGACGAGTTTCATGTTGTCGCCCGCTGCTCTCTTCCGGAGGGCTTCGCT 19409
QY 689 ACGCGCGGAGGTTCTACCGCATGAGTCCGTCATCAAGGAAAGGGCTGTCCACCG 748
Db 19408 GGGCGCGCGAGGTGTTCCACGCGCTGAAGAAGATTCTCAAGGGCGCAAGCTGGCCACGG 19349
QY 749 GACTTGGCGATGAGGGCGGCTTCGCTCTCTTCCGTCGCGCTCCACCGGTGAGGCTTTGACC 808
Db 19348 GCGTGGCGACGAGGGCGGTAGCCCGGACCTGCGCGGCAAGAGGCGCTGAAGC 19289
QY 809 TTATCGTTAAGGCAATCGAAGGCTGGCTTCAACCGAGGCAAGGATCGCTCTTGCTC 868
Db 19288 TCATCATGAGGCGCATCGACGAGCGGGCTTCAAGGCGCGTGAAGCGCTTCTCTGGGCC 19229
QY 869 TGGAGCTGTCTCTCTGAGTTCTTCAAGACGCGCACCTACCACTT-----CG 916
Db 19228 TGGACGTGGCGGCGAGGAGTTCTTCAAGAGGCGCACCTACCACTT-----CG 916
QY 917 AAGGTGGCCAGCACTCCGACGCTGAGATGGCAACGTTTACGCTGAGCTGTGTTGACCGGT 976
Db 19168 AGGGCAAGGATACGACTCCACCGGCTGTCTGAGTACTACCGGGGCTCTCCGAGCGCT 19109
QY 977 ACCCAATCTGCTCATCGAGGACCCCATCTGCAAGAGATGATGCGGAGGGTTACCAACAC 1036
Db 19108 ACCCCATCATCTCCATCGAAGCGGATGCGGAGGATGATGCGGAGGGCTGGAAGAAGC 19049

RESULT 7

US-09-252-991A-4449

; Sequence 4449, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; PRIOR FILING DATE: 1999-02-18

; PRIOR FILING DATE: 1998-02-18

; PRIOR FILING DATE: 1998-02-18

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; TYPE: DNA

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-4449

Query Match 33.7%; Score 532; DB 4; Length 1335;
Best Local Similarity 64.8%; Pred. No. 2.5e-139;
Matches 847; Conservative 0; Mismatches 440; Indels 21; Gaps 3;

QY 131 AGTTGCGATAGGAGGCGACAGTGGCTGGAATCATGACGCTATTGCTGCGGAAATTTCTCG 190
Db 26 ACTTCGGAGTGTAAACAAGAAATGGCAAGATCGTCGACATCAAGGGCGGTGAGTCTCTCG 85
QY 191 ACTCCCGCGGTAAACCAACCGTCGAGGAGAGGTTTTCTTGGATGACGTTTCCACCGGTG 250
Db 86 ACTCCCGCGGCAACCTTACCGTTGAAGCGGACGTGATCTCTGGACAACGGCATCGTCGGCA 145
QY 251 TCGCAGGTGTTCCATCCGCGCATCCACCGCGCTCCACGAGGCTCATAGCTCGGTGACG 310
Db 146 GCGCTTGGCGGCTTCCGCTGCTTCCACCGTTTCCCGGAGGCCCTCGAGCTGCGCATG 205
QY 311 GTGGCA---TCGCTACCTGGGCAAGGGCGTTTTGAAGGAGTTGAAAACGTCACCAAG 367
Db 206 GCACAAGAGCCGTTACCTGGGCAAGGGCGTGTGAAAGCCGTGGGCCAACATCAACGCC 265
QY 368 AAATCGGCGACGAGCTCGCTGGCTAGAGGCTGACGATCAGCGCTCATCGAAGCA 427

```
Db 266 CGATCCGCGACTGCTGCTGGGCAAGGACGCGCGGCGGACGAGAAAGCCCTCGACACCGCA 325
Qy 428 TGATCAAGCTTGATGGGACCGCAACAAAGTCCCGCTGGGTGCAAGCGCAATCTTGGTG 487
Db 326 TGATCGAGCTGGACGGCACCGAGAACAGGCAAGCTGGGGCCACGCGATCTTCGCGG 385
Qy 488 TTTCCATGGCTGTGCAAGGCTGCTGCTGATTCGCGAGGCTCCCACTGTTCGCTACA 547
Db 386 TGTCCCTGGCTGCGCCCAAGGCGCGCACAGGCAAGGGGTACCGCTGTACGCGACA 445
Qy 548 TCGGTGACCAACGCAAC-----ACGTTCTTCAGTTTCAATGATGAACATCA 595
Db 446 TCGCGGACCTCAACGCGACTCCCGGCGAGTACTCCATGCGCGTCCCGATGATGAACATCA 505
Qy 596 TCAACGCTGGCGCTCAACGCTGACTCCGCTGTTGAGTTTCAGGAATTCATGATCGCTCAA 655
Db 506 TCAACGCGCGGAGCATGCGGACAAACAGCTCGATATCCAGGAATTCATGTCAGCGCG 565
Qy 656 TCGGTGACAGACCTTCTCTGAGGCTCTCGGCAACGCGCGGAGGTCTACCAAGCACTGA 715
Db 566 TCGGCGGAGAACTTTCGCGAGGCGCTGCGCATGGGCGCGAGATCTTCATCACCTCA 625
Qy 716 AGTCCGCTCATCAAGGAAGGCTGCTGTCACCGGACTTTGGCGATGAGGGCGGCTTCGCTC 775
Db 626 AGGCGGTGCTGAAGGCGCGTGGCTGAACACCGCGTCTGCTGACGAAGGCGGCTTCGCGC 685
Qy 776 CTTCCGTCGCTCCACCGGTGAGGCTCTTGACCTTATCGTTAAGGCAATCGAAGGCTG 835
Db 686 CGAACCTGTGCTCAAGCAAGACGCTTGGCGGCGCATCGCGAGGCGCTCGAGAAGCGG 745
Qy 836 GCTTCAACCCAGGCAAGCAATCGCTCTGCTCTGGACGTGCTTCTCTGAGTTCTTCA 895
Db 746 GCTACAAGCTGGGCGACGAGTGAACCTTGGCCCTGGACTGGGCTTCAGCGAGTTCTTCA 805
Qy 896 AGGACGCGCACTACCACTTCGAAGGTGGCCAGCA-----CTCGCGAGCTGAGATGSCAA 949
Db 806 AGGACGCGCAAGTACGACTCGAAGCGCAAGCAAGGTATTTCGACGCGCGGTTTCGCGC 865
Qy 950 ACGTTTACGTGAGCTGTTGACGCGTACCCATCTGCTCCATCGAGGACCCACTGACG 1009
Db 866 ACTACCTGGCGCGGCTGACCCAGCGCTACCCGATCATCTCCATCGAGGACGCGATGACG 925
Qy 1010 AAGATGACTGGAGGGTTACACAACTCACCGCAACCATCGGCGCAAGGTTTCAGATCG 1069
Db 926 AGTCCGACTGGGCGGCTGGAAGGCTTGACCGACAGATCGGCGCCAAAGTCCAACTGG 985
Qy 1070 TTGGCGACGACTTCTTCGTCACCAACCTGAGCGCTGGAAGGAGGCGATCGCTAAGAAGG 1129
Db 986 TCGGCGACGACCTGTTCTGACCAACACCAAGATCTCTCAAGGAAGGATCGAGAAGGCA 1045
Qy 1130 CTGCGCAACTCATCTCTGTTAAGGTGAACAGATCGGTACCTCACCAGAGACTTCGACG 1189
Db 1046 TCGGCAATTCGATCTGATCAAGTTTCAACAGATCGGTTGCTCACCAGACCCCTGAGG 1105
Qy 1190 CTGTCGACATGGCTCACCGCGAGGCTACCTCCATGATGTCCACCGTTTCGCGTGAGA 1249
Db 1106 CCATCCAGATGGCCAGGCGCGCGCTATACCGGGTGATCTCGACCGCTTCGCGCGAAA 1165
Qy 1250 CCGAGGACACCATTTGCTGACCTCGCAGTTGCACTCAACTGTGGCGAGATCAAGACTG 1309
Db 1166 CCGAGGACTCGACCATCGCCGACCTTGGCGGTGAGTACCGCGCGGTCAGATCAAGACG 1225
Qy 1310 GTGCTCAGACGCTTCGACGCTGTCGAAAGTACAAACGAGTTCTCCGATCGAGCAGC 1369
Db 1226 GTTCGCTGTCGCGCTCGGACCGCGTGTCAAAGTACAAACCAAGTTGCTCGCATCGAAGC 1285
Qy 1370 TGCTTGGGACGCGCGGCTTACGAGGTGCGAGCGATTCCCAAGCT 1417
Db 1286 AACTGGGCGCAAGGCGCGCTACCGGTGTCGCGGGAATTCGCGGCT 1333
```

```
US-09-252-991A-4158/c
; Sequence 4158, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 4158
; LENGTH: 1347
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-4158
```

```
Query Match 33.7%; Score 531.4; DB 4; Length 1347;
Best Local Similarity 65.0%; Pred. No. 3.6e-139;
Matches 841; Conservative 0; Mismatches 431; Indels 21; Gaps 3;
Qy 146 CCACAGTGGCTGAAATCATGCAGTATTGCTCGCGAAATTCCTCGACTCCCGGGTAACC 205
Db 1343 CAAGATGGCAAGATCGTCGACATCAAGGCGCTGAGGTCTTGACTCCCGGGCAACC 1284
Qy 206 CAACCTCGAGGAGAGGTTTCTTGATGACGGTTCCTCCACGGTGTGCGAGGTGTTCCAT 265
Db 1283 CTACCGTTGAAGCGGAGCTGATCTTGACAAACGCGATCGTCGCGAGCGCTGCGCGCTT 1224
Qy 266 CCGGCGCATCCACGGGTTCCAGAGGCTCATGAGTGTGCTGAGGTGGGA---TCGCT 322
Db 1223 CCGGTCTTCCACCGTTCCCGGAGGCGCTCGAGTGTGCGGATGCGCAAGAGCGGTT 1164
Qy 323 ACCTGGGCAAGGCGGTTTGAAGGCAAGTTGAAACGTCACACGAAGAAATCGGCGACGAGC 382
Db 1163 ACCTGGGCAAGGCGGTTGCTGAAAGCGTGGCAACATCAACGCGCGATCCGCGACTGC 1104
Qy 383 TCCTGCGCTTAGAGGTGACGATCAGGCGCTCATTCGACGAAGCAATGATCAAGCTTGATG 442
Db 1103 TGCTGGGCAAGGACGCGGCGGACAGAAAGCCCTCGACACGCGATGATCGAGCTGGACG 1044
Qy 443 GCACCGCAACAAAGTCCCGCTGGGTGCAACCGCAATCTTGGTGTTCATGGCTGTTG 502
Db 1043 GCACCGAGAACAAAGGCGCAAGCTGGGCGCCAAACGCGATCTCTCGCGGTGTCTCGCTGCCG 984
Qy 503 CAAAGGCTGCTGCTGATTCGCGAGGCTCCCGCTGTTCCGCTACATCGGTGGACCAACG 562
Db 983 CCAAGGCGCGGACAGGCGCAAGGGGCTACCGGTGACGGCGACATCCCGACCTCAACG 924
Qy 563 CAC-----ACGTTCTTCCAGTTCCAAATGATGAACATCATCAACGCTGGCGCTC 610
Db 923 GCATCCCGGCCAGTACTCCATGCGGTGCGGATGATGAACATCATCAACGCGCGGAGC 864
Qy 611 AGCTGACTCCGGTGTGAGTTGAGTTGAGGAATTCATGATTCGCTCCAAATCGGTGCGAGACCT 670
Db 863 ATCGGACAAACAAAGTCTGATATCCAGGAATTCATGTTCCAGCGGTTCGCGCGGCAAGAACT 804
Qy 671 TCTCTGAGGCTCTCCGCAACGCGCGGAGGTCTACACGCACTGAGTCCGTCATCAAGG 730
Db 803 TCGCCAGGCGCTGCGCATGAGGCGCGGAGATCTTCCATCATCAAGGCGGCTGCTGAGG 744
Qy 731 AAAAGGCGCTGTCCACCGGACTTTGGCGATGAGGCGGCTTTCGCTCTCTTCGCTGGCTCCA 790
Db 743 CCGTGGCTTGAACAACGCGCTGCTGAGCAAGGCGGCTTCGCGCGGCAACCTGTCGCTCCA 684
Qy 791 CCGTGAAGCTCTTGACCTTATCGTTAAGCAATTCGAGAAAGGTGCTTCAACCCAGGCA 850
Db 683 ACGAAGACGCGCTGGCGCGCATCGCGAGGCGCGCTCGAGAAGGCGCGCTACAAGCTGGCGC 624
```


QY 851 AGGACATGCTTCTTGCTCTGAGCTTGCTTCTCTGAGTCTTCAAGGACGGCACCTACC 910
 DB |||||
 QY 623 ACGACGTGACCCCTGGCCCTGACTTGGCCCTCCAGGAGTCTTCAAGGACGGCAAGTACG 564
 DB |||||
 QY 911 ACTTGAAGGTGGCCAGCA-----CTCCGAGCTGAGATGGCAAAAGTTTACGCTGAGC 964
 DB |||||
 QY 563 ACCTGGAAGCGAAGGCAAGGATATTCGACGCGCCGGTTTCGCCGACTACTGCGCCGCC 504
 DB |||||
 QY 965 TCGTTGAGCGGTACCCCAATCGTCTCCATCGAGGACCCACATGCGAGGAAGATGACTGGGAGG 1024
 DB |||||
 QY 503 TGACCCAGCGCTACCCCGATCATCTCCATCGAGGACGGCATGGAGAGTCCGACTGGGCCG 444
 DB |||||
 QY 1025 GTTACACAAACCTCACCGCAACCATCGCGCAAGGTTTCAGATCGTTTGGGAGACTTCT 1084
 DB |||||
 QY 443 GCTGGAAGGCTGACCGACAGATCGCGCCCAAGGTCCAACTGGTCGGCGACGACCTGT 384
 DB |||||
 QY 1085 TCGTCAACAAACCTGAGCGCTGAAGGAGGCGATCGCTAAGAGGCTGCCAACTCCATCC 1144
 DB |||||
 QY 383 TCGTGACCAACACCAAGATCCTCAAGGAAGGCATCGAGAAGGGCATCGGCAATTCGATCC 324
 DB |||||
 QY 1145 TGGTTAAGGTGACCATGCTGCTACCTCACCAGACCTTCGAGCGTGTGACATGGCTC 1204
 DB |||||
 QY 323 TGATCAAGTTCAACACAGATCGGTTCGCTCAACGAGACCCCTGGAGGCCATCCAGATGGCCA 264
 DB |||||
 QY 1205 ACCGCGCAGGCTACACCTCTCATGATGTCCACACGTTCCGCTGAGACCGAGGACACCA 1264
 DB |||||
 QY 263 AGCGCGCGGTATACCGCGGTGATCTCGACCGCTCCGGGAAACCGAGGACTCGHACCA 204
 DB |||||
 QY 1265 TTGCTGACCTGGCAGTTGCACTCACTGTGTGCCAGATCAAGACTTGGTGTCCAGCAGTT 1324
 DB |||||
 QY 203 TCGCGACCTGGCGTGGGTACCGCCCGGTCAAGTCAAGACCGGTTCGCTGTGGCGCT 144
 DB |||||
 QY 1325 CCGACCGTTCGCAAGTACACAGCTTCTCCGCATCGAGCAGCTGCTTGGGACGCCG 1384
 DB |||||
 QY 143 CCGACCGCGTGTCAAGTACCAACAGTGTGTGGCGATCGAAGAGCAACTGGCGGCCAAG 84
 DB |||||
 QY 1385 GCGTCTACGAGGTGCGCAGCGCATTCACCGCT 1417
 DB |||||
 QY 83 CGCGTACCGTGTGCGCGGAATTCGCGGCT 51
 DB |||||

RESULT 9
 US-09-489-039A-2890
 ; Sequence 2890, Application US/09489039A
 ; Patent No. 6610836
 ; GENERAL INFORMATION:
 ; APPLICANT: Gary Breton et. al
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
 ; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 2709,2004001
 ; CURRENT APPLICATION NUMBER: US/09/489,039A
 ; CURRENT FILING DATE: 2000-01-27
 ; PRIOR APPLICATION NUMBER: US 60/117,747
 ; PRIOR FILING DATE: 1999-01-29
 ; NUMBER OF SEQ ID NOS: 14342
 ; SEQ ID NO 2890
 ; LENGTH: 1380
 ; TYPE: DNA
 ; ORGANISM: Klebsiella pneumoniae
 US-09-489-039A-2890

Query Match 30.4%; Score 479.6; DB 4; Length 1380;
 Best Local Similarity 62.5%; Pred. No. 1.4e-124;
 Matches 812; Conservative 0; Mismatches 464; Indels 24; Gaps 3;
 QY 140 AGAGGCCACAGTGGTGAATCATGCATATTCGCTCCGAAATTCGACTCCCGCG 199
 DB |||||
 QY 71 AGGAAACCTTAATGTCCAAAATCGTTAAAGTCATCGGTGTAATCATCGACTCCCGTG 130
 DB |||||
 QY 200 GTAACCCAAACCGTCGAGGACAGGTTTCTCGATGACGTTCCACGGTGTGCGAGGTG 259
 DB |||||
 QY 131 GTAACCCGACTGTTGAAGCCGAAGATACACCTGGAAGGTGTTTCGTCGGTATGGCAGCTG 190
 DB |||||

QY 260 TTCCATCCCGCGCATCCACCGGCGTCCACGAGGCTCATGAGCTGCGTGACGTTGGCGA-- 317
 DB |||||
 QY 191 CTCGCTCAGGTGCTTCTACTGGTTCCGCGCAAGCGTGGAACTGGCGGATGGGACAAAT 250
 DB |||||
 QY 318 -TCGCTACCTTGGCAAGGGGTTTTGAAGGACAGTTGAAACACGTTCAACGAAAGAAATCGCGC 376
 DB |||||
 QY 251 CCGGTTTCTGGGTAAGGCGTAACCAAGCTGTTGCTCGGTAAACGCGGCCGATCGCTC 310
 DB |||||
 QY 377 ACAGCTCGCTGCGCTAGAGGCTGACGATCAGCGCTTCATCGACGAAGCAATGATCAAGC 436
 DB |||||
 QY 311 AGCAATCTCTGGGCAAAAGATGCTAAAGATCAGGCTGGCATCGCAACAAGATCATGATGACC 370
 DB |||||
 QY 437 TTGATGGCACCGCCCAACAGTCCGCGTGGGTGCAACGCAATCTTGGTGTTCATG 496
 DB |||||
 QY 371 TGGACGGTACTGAAAAAAATACTAACTTCGGTTCGAAACGCTATCTCTGGCGCTTTTCTGG 430
 DB |||||
 QY 497 CTGTTGCAAGGCTGCTGCTGATTCGCGCAGGCGCTCCCACTGTTCCGCTACATCGTGGAC 556
 DB |||||
 QY 431 CGAACGCCAAGCAGCAGCTGCGTCCAAAGGCTTCGCGTGTAGCGGCATCGCTGAAC 490
 DB |||||
 QY 557 CAAACGG-----ACAGTTCTTCCAGTTCCAAATGATGAACATCATCAACGGTG 604
 DB |||||
 QY 491 TGAACGGCACTCCGCGCAAAATACTCCATGCGGTTCCGATGATGAACATCATCAACGGCG 550
 DB |||||
 QY 605 GCGCTCAGCGTGAATCGGTTGAGTTTCAGGAATTCATGATCGCTCCCAATCGGTGAC 664
 DB |||||
 QY 551 GTGAGCAGCTGACAAACAGCTCGATCCAGAAATTCATGATTCAGCGGTTGGGCGC 610
 DB |||||
 QY 665 AGACCTTCTCTGAGGCTCTCCGCAACGGCGGAGGTCTACCAACGCACTCAAGTCCGTC 724
 DB |||||
 QY 611 CGACCTGAAAGAGCAGTACGATCGGTTCTGAAGTTTCCATCACTGGCAAGTGC 670
 DB |||||
 QY 725 TCAAGAAAGAGGCTGTCCACCGGACTTGGCGATGAGGCGGCTGCTCTTCGTCG 784
 DB |||||
 QY 671 TGAAGTCCAAAGGCAACACATGCGGTTGGTACGAAAGGCGCTACGCGCCAACTGG 730
 DB |||||
 QY 785 GCTCCACCGTGAAGGCTTTCGCTTATCGTTAAGGCAATCGAAGAGGCTGGCTTACCC 844
 DB |||||
 QY 731 GTTCCAAAGCGGAGCGCTGGCTGTTATCGTGAAGCGTTAAAGTGCAGGCTACGAGC 790
 DB |||||
 QY 845 CAGGCAAGGACATCGCTTCTGCTCTGGAGCTTCTCTCTGAGTTCTTCAAGGACGGCA 904
 DB |||||
 QY 791 TGGGCAAGAGACATCACTTGGCGATGAGTCTGCGCGGCTCTGAATCTTCAAAAGACGTA 850
 DB |||||
 QY 905 CTTACCATTCGAAGTGGCGCAGCACT-----CCGAGCTGAGATGGCAAAAGCTTT 955
 DB |||||
 QY 851 AATACGTTCTGGCTGGCGAAGGCAACAAAGCGTTTCACTCTGAAGAGTTTCACTCACTTC 910
 DB |||||
 QY 956 ACGCTCAGCTCGTTGACGGGTACCCCAATCGTCTCCATCGAGGACCCACTGCGAGAGATG 1015
 DB |||||
 QY 911 TGGAAAGCTGACCAACACAGTACCCGATCGTCTCCATCGAAGATGGTCTGGACGAATCTG 970
 DB |||||
 QY 1016 ACTGGAGGTTTACCAACCTCACCGCAACCAATCGCGCAACAGTTTCAGATCGTTGGCG 1075
 DB |||||
 QY 971 ACTGGAAAGTTTTCGCTTACAGACTTAAAGTACTGGCGGACAAAAATCCAGCTGTTGGTG 1030
 DB |||||
 QY 1076 AGCACTTCTTCGTCACCAACCTTGAGGCGCTGAGGAGGCGCATCGCTAAGAAGGCTGCCA 1135
 DB |||||
 QY 1031 ACAGCTGTTTCGTAACCAACCAAGATCTTGAAGAAGGCAATCGAAAAAGGCAATCGCTA 1090
 DB |||||
 QY 1136 ACTCCATCTGGTTAAGGTGAACAGATCGGTATCCCTCACCGAGACCTTCGACGCTGTGCG 1195
 DB |||||
 QY 1091 ACTCCATCTGATCAAGTTCAACAGATCGGTTCTCTGACCGAAATCTTGGCTGCTATCA 1150
 DB |||||
 QY 1196 ACATGCTCACCGCGCAGGCTACCTCCATGATGTCCACCGTTCCGTTGAGACCGAGG 1255
 DB |||||
 QY 1151 AGATGGCAAGAGCGTGGCTACACCGCTGTTATCTCTCACCGTTCTGGCGAAACTGAAG 1210
 DB |||||
 QY 1256 ACACCAACCATTCGTGACCTCGAGTTGCACTCAACTGTGGCCAGATCAAGACTGGTGCTC 1315
 DB |||||
 QY 1211 ACCTACCATTCGCGACCTGGCTGTTGGTACCGCTGACGCGCAGATCAAAACTGGTTCTA 1270
 DB |||||
 QY 1316 CAGCAGTTTCCGACCGCTGTCGAAAGTACAAACCAAGCTTCTCCGCATCGAGCAGCTGCTT 1375
 DB |||||

Db 1271 TGAGCCGTTCTGACCGGTTGCTAAATACACACAGCTGATTCGTATCGAAGAAGCGCTGG 1330
Qy 1376 GCGAGCGCGGCTCTACGAGGTCGACGCGCATTCGCCAG 1415
Db 1331 GTGAGCAAGCGCGTTCAACGTCGTAAAGAGATCAAGG 1370

RESULT 10

US-09-818-780-22
; Sequence 22, Application US/09818780
; Patent No. 6677146
; GENERAL INFORMATION:
; APPLICANT: McHenry, Charles
; TITLE OF INVENTION: NOVEL THERMOPHILIC POLYMERASE III HOLOENZYME
; FILE REFERENCE: 1794.0030004
; CURRENT APPLICATION NUMBER: US/09/818,780
; CURRENT FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: US 60/192,736
; PRIOR FILING DATE: 2000-03-28
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 22
; LENGTH: 2363
; TYPE: DNA
; ORGANISM: Thermus thermophilus
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Includes 5' UTR and 3' UTR
US-09-818-780-22

Query Match 28.5%; Score 449.2; DB 4; Length 2363;
Best Local Similarity 64.3%; Pred. No. 6.6e-116;
Matches 692; Conservative 0; Mismatches 378; Indels 6; Gaps 1;

Qy 160 ATCATGACGATTCGCTCGGAAATTCCTGACTCCCGCGGTAAACCCACCGTCGAGCA 219
Db 1279 ATCTCGCGCTCCGGGACGCGAGGTTTGGATTCACAGGGGCTTTCCACGGTAGAGCG 1338
Qy 220 GAGGTTTTCTGGATGACGTTCCACGGTGTCCAGGTGTTCATCCGGGATCCACC 279
Db 1339 GAGGTGAGCTGGAAGCGGGGCGAGGGCGGCGCCATGTGGCCCTCCGGGGCTCCACC 1398
Qy 280 GGGCTCCACGAGGCTCATGAGCTCGGTGACGGTGGCGATCGTACCTGGGCAAGGGCGTT 339
Db 1399 GGAACCCACGAGGCTTGAGCTCAGGACGGCGGCAAGCGCTACCTGGGCAAGGGGTG 1458
Qy 340 TTGAAGCAGTTGAAACGTTCAACGAAGAATCGGCGACGAGCTCGTGGCCTAGAGGCT 399
Db 1459 CGCGGGCGGTGGAGAAGCTCAACGAGCGCATCGCCCGGAGCTCGTGGCATGGAGCC 1518
Qy 400 GACGATCAGCGCTCATCGAGAGCAATGATCAAGCTTGATGGCAGCGCCACAGTCC 459
Db 1519 CTGACAGGAAGGGGTGGAGCCGGCCATGCTGGAGTGGACGCGCACCCCAACAGGCC 1578
Qy 460 CGCTGGGTGCAACGCAATCTTGGTGTTCATGCTGTGCAAAAGGTGCTGCTGAT 519
Db 1579 AACCTGGAGGAACGCGCTCTCGCGTCTCCCTGGCGGTGGCCGGGGCGGCCGAG 1638
Qy 520 TCCGAGCGCTCCACATGTTTCCGTATCATCGTGGACCAACGACACGATTTCTTCCAGTT 579
Db 1639 GCCCTGGGCTGCCCTTTTACCGCTACTCTGGGGGGTTCAGGGGGTCACTTCCGCCGTG 1698
Qy 580 CCAATGATGACATCATCAACGGTGGCGCTCAGCTGACTCCCGTGTGAGGTTTCAGGAA 639
Db 1699 CCCCTCATGACGTCATCAACGGGGGAAGCAGCCGACCAACCGGGTGGACTTCCAGGAG 1758
Qy 640 TTCAATGCTCGCTCAATCGGTGAGAGACCTTCTCGAGGCTCTCCGCAAGCGCGGAG 699
Db 1759 TTTATGCTGGTCCCGGGGGGGAGAGCTTTCGCCGAGGCTTGAGGATCGGGGCCGAG 1818
Qy 700 GTCTACCAACGACTGAAGTCCGTATCAAGGAAAGGGCCCTGTCCACCGGACTTGGCGAT 759

Db 1819 GTCTTCCACACCTCAAGCGCGTCTCAAGAGAAGGGCTTACAGCACCAACAGTGGGGAC 1878
Qy 760 GAGCGGGCTTCGCTCTCTCCGTCGGCTCCACCGTGAGGCTCTTGACCTTATCGTTAAG 819
Db 1879 GAGGAGGCTTCGCCCGGACCTCAGAGCAACGAGAGCGGTGGAGCTTTTGCTCCTC 1938
Qy 820 GCAATCGAAGAGCTGCTTTCACCCAGCAAGGACATCGCTCTTGTCTGGAGCTTGTCT 879
Db 1939 GCCATTGAGCGGGGGGTACACCCCGGGCCAGAGGTCTCCCTGGCGCTTGACCCGGCC 1998
Qy 880 TCCTCTGAGTTCTTCAAGGAGCGCACTTACCACTTCGAAGGTG-----GCCAGCACTCC 933
Db 1999 ACAGCGAGCTTTACCGGGACGGGAAGTACCACTTGAAGGGGAGGCAAGGTCTCTCC 2058
Qy 934 GCAGCTGAGATGCAAAACGTTTACGCTGAGCTCGTTGACGCGTACCCCAATCGTCTCCATC 993
Db 2059 TCGAGGAGATGTTGGCTTCTCGGAGGCTCGGTGGAGAGTACCCCATCCGCTCCATT 2118
Qy 994 GAGGACCACTGCAAGAAAGTACTGGAGGGTTTACACCAACCTCACCGCAACCATCGGC 1053
Db 2119 GAGGACGGCTTCCGAGGAGCACTGGAGGGGTGGCGGCTTCTCACCGAGCGCTGGG 2178
Qy 1054 GACAAGTTTCAAGTCTGTCGAGCACTTCTTCTGTCACCAACCTTACGAGCGCTGAAAGAG 1113
Db 2179 GGAAGGTTCCAGCTCGTGGGGGACGACCTTCTGTCACCAACCGGAAAGGCTCCGGCG 2238
Qy 1114 GGCATCGCTAAGAAAGGCTGCCTTCCATCTCGTTAAGGTGAACAGATCGGTACCTC 1173
Db 2239 GGAATTGAGCGGGGGTGGCCAAACGCTTCTGTCAGGTGAACAGATCGGAGCCCTC 2298
Qy 1174 ACCGAGACCTTCGACGCTGTCGACATGGCTCACCGCGAGGCTACACCTCATGAT 1229
Db 2299 TCGGAACCTTCGAGGCACTCCCTGGCCAGCGCTCGGGGTACAGGGCGGTGAT 2354

RESULT 11

US-08-781-802-9
; Sequence 9, Application US/08781802
; Patent No. 5969121
; GENERAL INFORMATION:
; APPLICANT: ALLEN, Larry
; APPLICANT: AIKENS, John
; APPLICANT: FOSTEIN, Michael
; APPLICANT: VONSTEIN, Veronika
; APPLICANT: DEMIRJIAN, David
; APPLICANT: CASADABAN, Malcolm
; TITLE OF INVENTION: Stable Biocatalysts for Ester Hydrolysis
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff
; STREET: 300 S. Wacker Drive 32nd Floor
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781,802
; FILING DATE: 10-JAN-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/694,078
; FILING DATE: 07-AUG-1996
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/019,580
; FILING DATE: 12-JUN-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/009,704

/ FILING DATE: 11-JAN-1996
 / PRIOR APPLICATION DATA: US 60/001,995
 / FILING DATE: 01-AUG-1995
 / ATTORNEY/AGENT INFORMATION:
 / NAME: Chao, Mark
 / REGISTRATION NUMBER: 37,293
 / REFERENCE/DOCKET NUMBER: 95,963-E
 / TELECOMMUNICATION INFORMATION:
 / TELEPHONE: 312-913-0001
 / TELEFAX: 312-913-0002
 / INFORMATION FOR SEQ ID NO: 9:
 / SEQUENCE CHARACTERISTICS:
 / LENGTH: 3545 base pairs
 / TYPE: nucleic acid
 / STRANDEDNESS: single
 / TOPOLOGY: not relevant
 / MOLECULE TYPE: DNA (genomic)
 / ANTI-SENSE: NO
 / FEATURE:
 / NAME/KEY: CDS
 / LOCATION: 1397..2905
 / OTHER INFORMATION: /note= "E019 sequence of longest
 / OTHER INFORMATION: open reading frame; upstream untranslated region not exact"
 / FEATURE:
 / NAME/KEY: mat_peptide
 / LOCATION: 1397..2905
 / US-08-781-802-9

Query Match 27.9%; Score 440.8; DB 2; Length 3545;
 Best Local Similarity 62.7%; Pred. No. 1.9e-113;
 Matches 742; Conservative 0; Mismatches 427; Indels 15; Gaps 3;
 229 CTGATGACGGTCCCAACGGTGTGCGAGGTGTTCCATCCGCGCATCCACCGCGTCCAC 288
 39 CGGAAGAAGCGGTTTCGCGCGTGTAGTGCAAGCGCGGTTTCGACGGGGAATAT 98
 289 GAGCTCATGAGCTGCGTGAAGGTGC --GATCGCTACCTGGCAAGCGGCTTTGAAG 345
 99 GAAGCGGTTGAATGCGTGACGGCGCAAAACCGCTACCTCGGCAAGGGGTGCTCAA 158
 346 GCAGTTGAAACCGTCAACGAAGAAATCGGCGAGCTCGTGGCTTAGAGGCTGACGAT 405
 159 GCGGTTGAGAAGCTCAACGAAGTGATTGCTCCGGAATCATCGCTTAGAAGTACGTAT 218
 406 CAGCGCTCATCGACGAAGCAATGATCAAGTTGATGGCACCGCCAAACAAAGTCCCGCTG 465
 219 CAAGTGGCGATCGACCGCGGTTGATTGAATTTGACGCGACGGAACAAAGGAAAGCTT 278
 466 GGTGCAACCAATCCCTTGGTGTTCATCGCTGTGCAAGCTGCTGATTCGCA 525
 279 GGGCGAATGCTATTTAGGCGTGTGCTCGCGTGTGCTCGCGTGGGCTGATGAGCTT 338
 526 GGCTCCCACTGTTCCGCTATACGTGGTGAACCAACGACACAGTTCTTCCAGTTCCAATG 585
 339 GGCTTGGCTGTGACCAATACTTGGGCGGCTTTAAACGCTAAACGCTGCTGTACCGATG 398
 586 ATGAACATCATCAACGCTGGCGCTCAGCTGATCCGGTGTGACGTTTCAGGAATTCATG 645
 399 ATGAACATTTTAAACGCGCGCGCATGCGGACAAACAACTTGACATTTCAAGAAATTCATG 458
 646 ATCGCTCCAACTCGTGAGAGACTTCTGAGGCTCTCGCAACGCGCGGAGGTCTAC 705
 459 ATATGCGGCTCGTGGGAAAGCTTCCGTGAAGCGCTGCGCATGGTTCAGAAATTTTC 518
 706 CAGCACTGAAGTCCGCTCATCAAGGAAAGGGCTGTCCACCGACTTGGCGATGAGGGC 765
 519 CATAGCTTAAAGGTGTGTTTAAAGCGAAAGGCTTACAAACGCGCTGTGCGTGACGAAGGC 578
 766 GGCTTCGCTCTTCGCTCGGCTCCACCGCTGAGGCTCTTGACCTTATCGTTAAGGAATC 825
 579 GGATTTGCTCCGAACCTTAAATCGAAGAAAGCGCTGCAACGATCATTTGAAGCGATC 638

QY 826 GAGAAGCTGCTTACCCCGAGGAGGACATCGCTCTTGCTCTGAGCGTTGCTCTCTCT 885
 DB 639 GAAAAAGCCGCTTACAAACGAGGCAACAAAGTATGCTGCTATGACGTTGCTGCTCG 698
 QY 886 GAGTTCTTCAAG-----GAGGCACTTACCACTT-----CGAAGTGGCAGCACTCC 933
 DB 699 GAGCTGTACAAAGAGAGATGCAATATCATTTGGAAGGCGAAGCGCTGCTCAAAACA 758
 QY 934 GCAGCTGAGATGCAAAACGTTTACGCTGAGCTGCTGTTGACGCGTACCAATCGCTCCATC 993
 DB 759 TCAGAAGAAATGCTTGTGTTGATGAAGAGCTTGTGTCGAAATATCGATCATCTCGATC 818
 QY 994 GAGGACCACTGCAGGAAGATGATGCGGAGGGTTACCAACCTCACCGCAACCATCGGC 1053
 DB 819 GAAGACGAGCTTACGAAAAATGATGGAAGGCGCATAACTGCTTACTGAGCGCTTGGC 878
 QY 1054 GACAAGTTCAGATCGTTGCGGACGACTTCTGTCACCAACCTGAGCGCTGAAGGAG 1113
 DB 879 CACAAAGTGCAGCTCGTGGTGACGACTTGTGTTGTAACGAAACGAAAACTGGCGGAA 938
 QY 1114 GGCACTGCTAAGAAAGGCTGCCAATCCATCTCTGTTTAAAGTGAACAGATCGGTACCTC 1173
 DB 939 GGCATTGAAAAAGCGCTGCGCAACTCGATTTTAAATTAAGTGAACCAATCGGTACACTG 998
 QY 1174 ACCGAGACCTTGCACGCTGTCGACATGGCTCACCGCGAGGCTACACCTCCATGATGTC 1233
 DB 999 ACGAAACGTTTCGATGCTTGCATGAGTGGCAAAACGCGCGGCTTACACGCGGTTGTGTC 1058
 QY 1234 CACGTTTCCGCTGAGACCGAGGACACCACTTGTGACCTCGCAGTTGCACTCAACTGT 1293
 DB 1059 CACGTTTCCGTTGAACGGAAGACACGATGTTGCCGATATCGCTGCGCAAAACGCT 1118
 QY 1294 GGCAAGTCAAGACTGCTGCTCCAGCACGTTCCGACCGTGTGCAAGTGTACAAAGTAC 1353
 DB 1119 GGCCTAAATCAAAACGCGGAGCACCGTTCGCTGCGGACCGGCTGCGCAAAATACAC 1178
 QY 1354 CTCGCTGAGCAGCTGCTTGGCGAGCGCGGCTCTACGAGG 1397
 DB 1179 CTCGCTGAGCAGCTTGGCGACGAACTTGGCCACGCGCTATTTACCAAGG 1222

RESULT 12
 US-09-583-110-1156
 ; Sequence 1156, Application US/09583110
 ; Patent No. 6699703
 ; GENERAL INFORMATION:
 ; APPLICANT: Lynn Doucette-Stamm et al.
 ; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
 ; TITLE OF INVENTION: Pneumoniae for Diagnostics and Therapeutics
 ; FILE REFERENCE: PATH00-07A
 ; CURRENT APPLICATION NUMBER: US/09/583,110
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: US 09/107,433
 ; PRIOR FILING DATE: 1998-06-30
 ; PRIOR APPLICATION NUMBER: US 60/085,131
 ; PRIOR FILING DATE: 1998-05-12
 ; PRIOR APPLICATION NUMBER: US 60/051,553
 ; PRIOR FILING DATE: 1997-07-02
 ; NUMBER OF SEQ ID NOS: 5322
 ; SEQ ID NO 1156
 ; LENGTH: 1305
 ; TYPE: DNA
 ; ORGANISM: Streptococcus pneumoniae
 US-09-583-110-1156

Query Match 27.4%; Score 431.6; DB 4; Length 1305;
 Best Local Similarity 59.8%; Pred. No. 4.3e-111;
 Matches 775; Conservative 0; Mismatches 494; Indels 27; Gaps 2;
 QY 160 ATCATGACGATATTCGCTCGGAAATTCGACTCCCGGGTAAACCAACCGTCGAGGCA 219
 DB 10 ATTACTGATGTTTACGCTCGGAAGTCTTAGACTCAGCGGTAACCCAACTTGAAGTA 69

Qy	220	GAGGTTTTCTGGATGACGGTTCCACGGTGTCCGAGGTGTTCCATCCGGCGCATCCACC	279
Db	70	GAAGTTTACACTGAATCAGGTGCTTCCGACGCTGATGGTTTCATCAGGAGCTTCTACT	129
Qy	280	GGCGTCCACGAGGCTCATGAGTCGCTGACGGTGGCGA---TCGCTACCTGGGCAAGGGC	336
Db	130	GGTGAAACACGAGCAGTTGAACCTTCGGACGGTGACAAATCTCGTTACGGTGTCTTGGT	189
Qy	337	GTTTTGAAGGCAGTTGAAAAACGTCAACGAAGAAATCGGCGACGAGCTCGCTGGCGCTAGAG	396
Db	190	ACAAAAAGCTGTTGACAAACGTAACCAATCATTTGCTGAAGCTATCATTTGGCTACGAT	249
Qy	397	GCTGACGATCAGCCCTCATCGACGAAGCAATGATCAAGCTTGATGGCACCGCCAAACAG	456
Db	250	GTACGTGATCAACAAAGCTATTGACCGTGCTATGATCGCACTTGACGAGTACTCTCTAAACAA	309
Qy	457	TCCCGCTGGGTGCAAAACGCAATCCTTTGGTGTTCATGGTGTTGCAAAAGGCTGCTGCT	516
Db	310	GGTAAATTGGGTGCGAATGCAATCCTCGGTGTGTCTATCGCTGTAGCTCGTGTGCTGCT	369
Qy	517	GATTCGCGAGGCTCCCACTGTTCCGCTACATCGGTGGACAAACGCAACAGTTCCTTCCA	576
Db	370	GACTACCTTGAAATCCCACCTTTACAGCTACCTTGCTGGATTCACACATAAAGTTCTTCCA	429
Qy	577	GTTCCAATGATGAACATCATCAACGGTGGCGCTCACGCTGACTCCGGTGTGAGCTTCAG	636
Db	430	ACTCCAATGATGAACATCATCAACGGTGGTTCCTCACTCGACGCTCCAAATCGCTTCCAA	489
Qy	637	GAATTCAATGATCGCTCCAAATCGGTGCAGAGACCTTCTCTGAGGCTCTCCGCAACGGCGC	696
Db	490	GAGTTCAATGATCTTGCCAGTTGGTGGCCAAACATTTAAAGAGGCCCTTCGTACGGTGCT	549
Qy	697	GAGGTCTACACGCACTGAAGTCGCTCATCAGAGAAAGGCCGTGTCCACCGCACTTGGC	756
Db	550	GAATCTTCCACGCGCTTAAGAAAATCCTTAAATCACGTGTTTTGGAAACTCGCGCTAGGT	609
Qy	757	GATGAGGCGGCTTCGCTCCTTCGTCGGCTCCACCGCTGAGGCTCTTGACCTTATCGTT	816
Db	610	GACGAAGTGATTCGCTCCTTCGTTTCCAGGAACCTGAAGATGGTGTGAAACTATCCCT	669
Qy	817	AAGSCAATCAGAAAGGCTGGCTTCAACCCGAGGCAAGGACATCGCTCTTGCTCTGGACGTT	876
Db	670	GCTCGAATGAAGCTGCTGGATATGTACCAGGTAAGACGCTATTTATTGGATTGACTGT	729
Qy	877	GCTTCTCTGAGTTCCTCAAGGACGGCACCTACCACTTCGA-----	917
Db	730	GCTTCAATCAGAAATTTCTACGATAAAGAACGTAAGTTTACGACACATAAAATTTGAAGGT	789
Qy	918	-----AGTGGCCAGCACTCCGACGCTGAGATGSCAAACGTTTACGCTGAGCTCGTTGAC	972
Db	790	GAAGTGCTGCTGTTGTTACATCTGCAAGAACAAATCGACTACCTTGAAGAAATTTGGTTAAC	849
Qy	973	CGGTACCCAATCGTCTCCATCGAGGACCCCATCGCAGGAAGATGACTGGGAGGGTTTACACC	1032
Db	850	AAATACCCAATCATCATATTGAAGATGGTATGGATGAAACGACTGGGATGGTTGGAAA	909
Qy	1033	AACTCACCGCAACCATCGGCGACAAAGTTTCAGATCGTTGGCGACGACTCTTTCGTCACC	1092
Db	910	GCTCTTACTGAAACGCTTGTGTAAGAAAGTACAACTTGTGTGGTACGACTCTTTCGTTAA	969
Qy	1093	AACCTCAGCGCCTGAAGGAGGGCATCGCTAAGAAAGGCTGCCAACTCCATCTCGTTTAA	1152
Db	970	AACACTGACTTACCTTGCACGTGGTATCCAAGAGGTGCTGCTAACTCAATCTTTATCCAA	1029
Qy	1153	GTGAACCCAGATCGGTACCCCTCACCGAGACCTTCAGACCTGTTCGACATGGCTCACCGCGCA	1212
Db	1030	GTTAAACCAATCGGTACTCTTACTTGAAACTTTTGAAGCTATCGAAATGGCTAAAGAGCT	1089
Qy	1213	GGGTACACCTCCATGATGTCACCGCTTCGGTGAGACCGAGGACCAACCATTCGCTGAC	1272
Db	1090	GGTTACACTGCTGTTGTTATCACACCGTTTCAGGTGAAATCTGAAGATTCACAAATCGCTGAT	1149

Qy	1273	CTCGCAGTTGCAC	TCAACTGTCGCACAGATCAAGACTGGTGTCTCCAGCAGCTTCCGACCGT	1332
Db	1150	ATTGCGAGTTTGC	AACTAACACGACGACAAATCAAGACTGGTTCACTTTCACGTACAGACGC	1209
Qy	1333	GTCCGAAAGTACA	CAACACAGCTTCTTCGCACTCGAGCAGCTGCTTTGGCGACGCCGGCTCTAC	1392
Db	1210	ATCGCTAAATACA	CAACCAATTTGTTTGTATCGAAGACCAACTTTGGTGAAGTAGCTGAATAT	1269
Qy	1393	GCAGTTCGCGAG	CGCATTTCCACGCTTTTTCAGGGCTAA	1428
Db	1270	CGTGGATTGAAT	CAATCTTACACACTTTAAAAAATAA	1305
RESULT 13				
US-08-961-527-45/c				
; Sequence 45, Application US/08961527				
; Patent No. 6420135				
; GENERAL INFORMATION:				
; APPLICANT: Charles Kunsch				
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences				
; NUMBER OF SEQUENCES: 391				
; CORRESPONDENCE ADDRESS:				
; ADDRESSEE: Human Genome Sciences, Inc.				
; STREET: 9410 Key West Avenue				
; CITY: Rockville				
; STATE: Maryland				
; COUNTRY: USA				
; ZIP: 20850				
; COMPUTER READABLE FORM:				
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage				
; COMPUTER: HP Vectra 486/33				
; OPERATING SYSTEM: MSDOS version 6.2				
; SOFTWARE: ASCII Text				
; CURRENT APPLICATION DATA:				
; APPLICATION NUMBER: US/08/961,527				
; FILING DATE:				
; CLASSIFICATION: 424				
; PRIOR APPLICATION DATA:				
; APPLICATION NUMBER:				
; FILING DATE:				
; ATTORNEY/AGENT INFORMATION:				
; NAME: Brookes, A. Anders				
; REGISTRATION NUMBER: 36,373				
; REFERENCE/DOCKET NUMBER: PB340P1				
; TELECOMMUNICATION INFORMATION:				
; TELEPHONE: (301) 309-8504				
; TELEFAX: (301) 309-8512				
; INFORMATION FOR SEQ ID NO: 45:				
; SEQUENCE CHARACTERISTICS:				
; LENGTH: 11384 base pairs				
; TYPE: nucleic acid				
; STRANDEDNESS: double				
; TOPOLOGY: linear				
US-08-961-527-45				
Query Match 27.3%; Score 430.4; DB 3; Length 11384;				
Best Local Similarity 59.4%; Pred. No. 2.9e-110;				
Matches 781; Conservative 0; Mismatches 506; Indels 27; Gaps 2;				
Qy	142	GAGGCCACAGTGC	TGAATCATGACGATTTTCGTCGCGAAATTTCTCGACTCCCGCGGT	201
Db	3064	GAGTAAAAAATG	CAATTATTACTGATGTTTACGCTCGCGAAGTCTTAGACTCAGCGGT	3005
Qy	202	AACCCAACCGT	CGAGCAGAGGTTTTTCTGGATGACGGTTCCACGGTGTCCGAGGTGTT	261
Db	3004	AACCCAACACTT	GAAAGTAGAAGTTTACACTGAATCAGGTGCTTTTCGGACGGTATGGTT	2945
Qy	262	CCATCCGGCGCAT	TCCACCGCGCTCCACGAGGCTCATGAGTCGCTGACGGTGGCGA--T	318
Db	2944	CCATCAGGAGCT	TCTACTGCTGAACACGACGAGTTTGAACTTCCGACCGGTGACAAATCT	2885
Qy	319	CGCTACTCTCGG	CAAGGGCGTTTTTTGAAGGCAGTTTGAHAACGTCACAGAGAAATCGGCGAC	378

Qy	1081	TTCTTCGTACCAACCTGAGCGCTGAACGAGGGCATCGCTAAGAGGCTGCCAACTCC	1140
Db	393	CTGTTCGTAAACCAACACCAAGATCCTGAAAAGAGGCATCGATAACTCC	334
Qy	1141	ATCTGGTTAAGGTGAACCAAGATCGGTACCTCACCGAGACCTTCGACGCTGTCGACATG	1200
Db	333	ATCTGATCAAGTTCAACCAAGATCGGTCTCTGACCGAAACTCTGGCTGCTATCAAGATG	274
Qy	1201	GCTCACCGCGCAGGCTACACCTCCATGATGCCACCGTTCCGGTGAGACCGAGGACACC	1260
Db	273	GCGAAGACGCTGGCTACACCGCTGTTATCTCTCACCGTTCTGGCGAAACTGAAGACGCT	214
Qy	1261	ACCATTCCTGACCTCGCAGTTGCACTCAACTGTGGCCAGATCAAGACTGGTGTCCAGCA	1320
Db	213	ACCATTCGCCACCTGGCTGTGTGTACCGCTGCGAGGCCAGATCAAAACTGGTTCATGAGC	154
Qy	1321	CGTTCCGACCGTGTGCAAAAGTACAACCAAGTCTTCGCGATCGAGCAGCTGCTTGGCGAC	1380
Db	153	CGTTCGACCGCGTTGCTAATACCAACCAAGCTGATTCGTATCGAAGAGCGCTGGGTGAG	94
Qy	1381	GCCGGCGTCTACGAGGTCGACGCGCATTCGCCAG	1415
Db	93	CAAGCGCGTTCAACGTCGTAAAGAGATCAAAGG	59

Search completed: September 28, 2005, 23:33:54
Job time : 293.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 28, 2005, 20:45:32 ; Search time 4241.5 Seconds
(without alignments)
2568.796 Million cell updates/sec

Title: US-10-728-947-3

Perfect score: 1578

Sequence: 1 ggctgggatatgggtagtt.....ctcaagcagggaactgtctt 1578

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 7442561 seqs, 3452328358 residues

Total number of hits satisfying chosen parameters: 14885122

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:*
- 18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq:*
- 19: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq:*
- 20: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq:*
- 21: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq:*
- 22: /cgn2_6/ptodata/2/pubpna/US10J_NEW_PUB.seq:*
- 23: /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq:*
- 24: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
- 25: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 26: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1578	100.0	1578	9	US-09-860-768-3
2	1578	100.0	1578	20	US-10-728-947-3
3	1576.4	99.9	1578	9	US-09-860-768-1
4	1576.4	99.9	1578	20	US-10-728-947-1
5	1576.4	99.9	3309400	9	US-09-738-626-1
6	1401.8	88.8	1405	21	US-10-494-836-7
7	1394.8	88.4	1398	19	US-10-781-014-71

8	1273.4	80.7	1275	9	US-09-738-626-1085	Sequence 1085, Ap	
9	907	57.5	1275	17	US-10-282-122A-17724	Sequence 17724, A	
c	10	696.8	44.2	9025608	15	US-10-156-761-1	Sequence 1, Appli
11	695.4	44.1	1284	15	US-10-156-761-3520	Sequence 3520, Ap	
12	678.2	43.0	1269	17	US-10-369-493-32027	Sequence 32027, A	
13	647	41.0	1287	17	US-10-282-122A-25949	Sequence 25949, A	
14	632.6	40.1	45191	15	US-10-080-170-649	Sequence 649, App	
15	632.6	40.1	45191	19	US-10-080-170-649	Sequence 649, App	
16	632.6	40.1	45191	19	US-10-468-356-649	Sequence 649, App	
17	632.6	40.1	1317	17	US-10-282-122A-26512	Sequence 26512, A	
18	626.2	39.7	1290	17	US-10-282-122A-28306	Sequence 28306, A	
19	583.4	37.0	1344	17	US-10-282-122A-27498	Sequence 27498, A	
c	20	583	2256646	19	US-10-470-565-1	Sequence 1, Appli	
21	580.4	36.8	1269	17	US-10-369-493-34244	Sequence 34244, A	
22	577	36.6	1269	17	US-10-369-493-35383	Sequence 35383, A	
23	575.6	36.5	1281	17	US-10-369-493-40598	Sequence 40598, A	
24	572.4	36.3	1275	17	US-10-369-493-35506	Sequence 35506, A	
25	569	36.1	1260	17	US-10-369-493-38426	Sequence 38426, A	
26	569	36.1	1269	17	US-10-369-493-38909	Sequence 38909, A	
27	568.2	36.0	1278	17	US-10-369-493-41573	Sequence 41573, A	
28	567.4	36.0	1272	17	US-10-369-493-44349	Sequence 44349, A	
29	561.8	35.6	1272	17	US-10-369-493-31449	Sequence 31449, A	
30	558.6	35.4	1278	17	US-10-369-493-45417	Sequence 45417, A	
31	550.4	34.9	1290	17	US-10-369-493-43178	Sequence 43178, A	
32	542	34.3	1281	17	US-10-282-122A-11620	Sequence 11620, A	
33	530.2	33.6	1290	9	US-09-815-242-7834	Sequence 7834, Ap	
34	530.2	33.6	1290	17	US-10-282-122A-30326	Sequence 30326, A	
35	529.2	33.5	1281	17	US-10-369-493-44693	Sequence 44693, A	
36	528	33.5	1287	17	US-10-369-493-32181	Sequence 32181, A	
37	522.8	33.1	1278	17	US-10-369-493-28375	Sequence 28375, A	
38	522.8	33.1	1281	17	US-10-369-493-31134	Sequence 31134, A	
39	522.8	33.1	1281	17	US-10-282-122A-13098	Sequence 13098, A	
40	513.6	32.5	1404	17	US-10-282-122A-33598	Sequence 33598, A	
41	513.2	32.5	1269	17	US-10-369-493-47120	Sequence 47120, A	
42	512	32.4	1293	17	US-10-369-493-46986	Sequence 46986, A	
43	507.8	32.2	1293	17	US-10-369-493-37536	Sequence 37536, A	
44	506	32.1	1293	9	US-09-974-300-664	Sequence 664, App	
45	503.2	31.9	1293	17	US-10-282-122A-9983	Sequence 9983, Ap	

ALIGNMENTS

RESULT 1

US-09-860-768-3
; Sequence 3, Application US/09860768
; Patent No. US20020082403A1
; GENERAL INFORMATION:
; APPLICANT: Mockel, Bettina
; APPLICANT: Pfefferle, Walter
; APPLICANT: Hermann, Thomas
; APPLICANT: Pohler, Alfred
; APPLICANT: Kalinowski, Jorn
; APPLICANT: Bathe, Brigitte
; TITLE OF INVENTION: New Nucleotide Sequences that Code for the Eno Gene
; FILE REFERENCE: 21123/278404
; CURRENT APPLICATION NUMBER: US/09/860,768
; CURRENT FILING DATE: 2001-05-21
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 3
; LENGTH: 1578
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
US-09-860-768-3

Query Match 100.0%; Score 1578; DB 9; Length 1578;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1578; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGCTGGGGATATGGGTAGTTTTCGCCCACTAAATTCACCTGATTCATCGAAACAAGA 60

Db 1 GGCTGGGGATATGGGTAGTTTTCGCCCACTAAATTCACCTGATTCATCGAAACAAGA 60

QY 61 TTCTGCAACAATTGGGTAGACGTGATTGAAGACATTTTGATCAGTGAATAATTCTAG 120
 Db 61 TTCTGCAACAATTGGGTAGACGTGATTGAAGACATTTTGATCAGTGAATAATTCTAG 120
 QY 121 TTAGCTCCCAAGTTGGCATAGGAGGCCACAGTGGCTGAATCATGCACGTATTCGCTCGC 180
 Db 121 TTAGCTCCCAAGTTGGCATAGGAGGCCACAGTGGCTGAATCATGCACGTATTCGCTCGC 180
 QY 181 GAAATTCGACTCCCGCGGTAAACCAACCGTGCAGGCAAGGTTTTCCTGGATGACGGT 240
 Db 181 GAAATTCGACTCCCGCGGTAAACCAACCGTGCAGGCAAGGTTTTCCTGGATGACGGT 240
 QY 241 TCCACGGTTCGACAGTGTTCATCCGCGGCATCCACCGCGCTCCACAGAGGCTCATGAG 300
 Db 241 TCCACGGTTCGACAGTGTTCATCCGCGGCATCCACCGCGCTCCACAGAGGCTCATGAG 300
 QY 301 CTGCTGACGGTGGCGATCGCTACCTGGGCAAGGGCGTTTGAAGGCAGTTGAAAACGTC 360
 Db 301 CTGCTGACGGTGGCGATCGCTACCTGGGCAAGGGCGTTTGAAGGCAGTTGAAAACGTC 360
 QY 361 AACGAAGAAATCGCGACGAGCTCGCTGGCTTAGAGGCTGACGATCAGCGCCTCATCGAC 420
 Db 361 AACGAAGAAATCGCGACGAGCTCGCTGGCTTAGAGGCTGACGATCAGCGCCTCATCGAC 420
 QY 421 GAAGCAATGATCAAGCTTGATGGCACCACCACCAAGTCCCGCTGGGTGCAAAACGCAATC 480
 Db 421 GAAGCAATGATCAAGCTTGATGGCACCACCACCAAGTCCCGCTGGGTGCAAAACGCAATC 480
 QY 481 CTTGGTGTTCCTATGGCTGTTCGAAAGGCTGCTGCTGATTCGCGAGGCTCCCACTGTTC 540
 Db 481 CTTGGTGTTCCTATGGCTGTTCGAAAGGCTGCTGCTGATTCGCGAGGCTCCCACTGTTC 540
 QY 541 CGCTACATCGGTGGAACAAAGGCACACGTTCTTCAGTTCCAATGATGAACATCATCAAC 600
 Db 541 CGCTACATCGGTGGAACAAAGGCACACGTTCTTCAGTTCCAATGATGAACATCATCAAC 600
 QY 601 GGTGGGCTCAGCTGACTCCGGTGTGAGTTGAGTTTCAAGAAATTCATGATCGCTCCAATCGGT 660
 Db 601 GGTGGGCTCAGCTGACTCCGGTGTGAGTTTCAAGAAATTCATGATCGCTCCAATCGGT 660
 QY 661 GCAGAGACCTTCTCTGAGGCTCTCCGCAACCGCGCGAGGCTTACCAACGCACTGAAGTCC 720
 Db 661 GCAGAGACCTTCTCTGAGGCTCTCCGCAACCGCGCGAGGCTTACCAACGCACTGAAGTCC 720
 QY 721 GTCATCAAGAAAGGCGCTGTCCACCGGACTTGGCGATGAGGGCGGCTTCGCTCCTTCC 780
 Db 721 GTCATCAAGAAAGGCGCTGTCCACCGGACTTGGCGATGAGGGCGGCTTCGCTCCTTCC 780
 QY 781 GTGGCTCCACCGGTGAGGCTTTGACCTTATCGTTAAGCAATCGAAGAGGCTGGCTTC 840
 Db 781 GTGGCTCCACCGGTGAGGCTTTGACCTTATCGTTAAGCAATCGAAGAGGCTGGCTTC 840
 QY 841 ACCCAGGCAAGGACATCGCTCTTGCTCTGAGAGTTCCTCTGAGTTCTTTCAAGGAC 900
 Db 841 ACCCAGGCAAGGACATCGCTCTTGCTCTGAGAGTTCCTCTGAGTTCTTTCAAGGAC 900
 QY 901 GGCACCTACCACTCGAAGGTGGCAGCACTCCGACGCTGAGATGGCAAAAGTTTACGCT 960
 Db 901 GGCACCTACCACTCGAAGGTGGCAGCACTCCGACGCTGAGATGGCAAAAGTTTACGCT 960
 QY 961 GAGCTCGTTGACGGTACCCTGCTGCTGAGAGTTCCTCTGAGTTCTTTCAAGGAC 1020
 Db 961 GAGCTCGTTGACGGTACCCTGCTGCTGAGAGTTCCTCTGAGTTCTTTCAAGGAC 1020
 QY 1021 GAGGTTTACCAACCTTACCGCAACCATCGGCAAGGTTTCAAGTTCGTTGGCGAGC 1080
 Db 1021 GAGGTTTACCAACCTTACCGCAACCATCGGCAAGGTTTCAAGTTCGTTGGCGAGC 1080
 QY 1081 TTCTTCGTCACCAACCTTACCGCAACCATCGGCAAGGTTTCAAGTTCGTTGGCGAGC 1140
 Db 1081 TTCTTCGTCACCAACCTTACCGCAACCATCGGCAAGGTTTCAAGTTCGTTGGCGAGC 1140

QY 1141 ATCTCGTTAAGGTGAACCCAGATCGGTACCTCACCGAGACCTTTCGACGCTGTCGACATG 1200
 Db 1141 ATCTCGTTAAGGTGAACCCAGATCGGTACCTCACCGAGACCTTTCGACGCTGTCGACATG 1200
 QY 1201 GCTCACCGCGCAGGCTACACCTCCATGATGTCCCACCGTTCGCGTGAGACCCGAGGACCC 1260
 Db 1201 GCTCACCGCGCAGGCTACACCTCCATGATGTCCCACCGTTCGCGTGAGACCCGAGGACCC 1260
 QY 1261 ACCATTGCTGACCTCCAGTTGACCTCAACTGTGGCAGATCAAGACTGGTGTCCAGCA 1320
 Db 1261 ACCATTGCTGACCTCCAGTTGACCTCAACTGTGGCAGATCAAGACTGGTGTCCAGCA 1320
 QY 1321 CGTTCCGACCGTGTCCGAAAGTACAAACAGCTTCCCGCATCGAGCAGCTGCTTGGCGAC 1380
 Db 1321 CGTTCCGACCGTGTCCGAAAGTACAAACAGCTTCCCGCATCGAGCAGCTGCTTGGCGAC 1380
 QY 1381 GCGGCGCTCTACGACGCTCGACGCAATTCACACGCTTTCAGGGCTTAAATAAAGCGCTT 1440
 Db 1381 GCGGCGCTCTACGACGCTCGACGCAATTCACACGCTTTCAGGGCTTAAATAAAGCGCTT 1440
 QY 1441 TTCGACGCGCGGTAACTCAAGTTCCCGGGCGTGGTTCGCTTACTACTGTTACTGTTGT 1500
 Db 1441 TTCGACGCGCGGTAACTCAAGTTCCCGGGCGTGGTTCGCTTACTACTGTTACTGTTGT 1500
 QY 1501 GACTATGATCGAGGATTATGGCAAGCAGAAAGAACTCATAAAGGCTTGTTCCTGTCT 1560
 Db 1501 GACTATGATCGAGGATTATGGCAAGCAGAAAGAACTCATAAAGGCTTGTTCCTGTCT 1560
 QY 1561 CAAGCAGGGAACGTGCTT 1578
 Db 1561 CAAGCAGGGAACGTGCTT 1578

RESULT 2
 US-10-728-947-3
 ; Sequence 3, Application US/10728947
 ; Publication No. US20040220394A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Mockel, Bettina
 ; APPLICANT: Pfefferle, Walter
 ; APPLICANT: Hermann, Thomas
 ; APPLICANT: Pohler, Alfred
 ; APPLICANT: Kalinowski, Jorn
 ; APPLICANT: Bathe, Brigitte
 ; TITLE OF INVENTION: New Nucleotide Sequences that Code for the Eno Gene
 ; FILE REFERENCE: 21123/278404
 ; CURRENT APPLICATION NUMBER: US/10/728,947
 ; CURRENT FILING DATE: 2003-12-08
 ; PRIOR APPLICATION NUMBER: US/09/860,768
 ; PRIOR FILING DATE: 2001-05-21
 ; NUMBER OF SEQ ID NOS: 6
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 3
 ; LENGTH: 1578
 ; TYPE: DNA
 ; ORGANISM: Corynebacterium glutamicum
 US-10-728-947-3

Query Match 100.0%; Score 1578; DB 20; Length 1578;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1578; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCGTGGGATATGGGTAGTTTTCGCCACATAATTTCACTGATTCGCTCATCGAAACAAGA 60
 Db 1 GCGTGGGATATGGGTAGTTTTCGCCACATAATTTCACTGATTCGCTCATCGAAACAAGA 60
 QY 61 TTCTGTCACCAATTTGGGTGTAGACGTGATTGAAGACATTTTGATCAGTGAATAATTCTAG 120
 Db 61 TTCTGTCACCAATTTGGGTGTAGACGTGATTGAAGACATTTTGATCAGTGAATAATTCTAG 120
 QY 121 TTAGCTCCCAAGTTGGCATAGGAGGCCACAGTGGCTGAATCATGCACGTATTCGCTCGC 180
 Db 121 TTAGCTCCCAAGTTGGCATAGGAGGCCACAGTGGCTGAATCATGCACGTATTCGCTCGC 180
 QY 121 TTAGCTCCCAAGTTGGCATAGGAGGCCACAGTGGCTGAATCATGCACGTATTCGCTCGC 180
 Db 121 TTAGCTCCCAAGTTGGCATAGGAGGCCACAGTGGCTGAATCATGCACGTATTCGCTCGC 180

181 GAAATTTCTGACTCCCGCGGTAAACCAACCGTTCGAGGCAGAGGTTTTCTTGGATGACGGT 240
 181 GAAATTTCTGACTCCCGCGGTAAACCAACCGTTCGAGGCAGAGGTTTTCTTGGATGACGGT 240
 241 TCCACGGTTCGAGGTGTTTCATCCGGGCGATCCACCGGCGTCCACGAGGCTCATGAG 300
 241 TCCACGGTTCGAGGTGTTTCATCCGGGCGATCCACCGGCGTCCACGAGGCTCATGAG 300
 301 CTGGTTCAGCGTTCGAGTTCGCTACCTCGGCAAGGGCGTTTTGAAGGCGATTGAAGAGTC 360
 301 CTGGTTCAGCGTTCGAGTTCGCTACCTCGGCAAGGGCGTTTTGAAGGCGATTGAAGAGTC 360
 361 AACGAGAAATTCGCGGACGAGCTCGCTGGCTAGAGGCTGACGATCAGCGCCTCATCGAC 420
 361 AACGAGAAATTCGCGGACGAGCTCGCTGGCTAGAGGCTGACGATCAGCGCCTCATCGAC 420
 421 GAAGCAATGATCAAGCTTTCGATGGCAGCGCAACAAAGTCCCGGCTGGGTGCAAGCGCAATC 480
 421 GAAGCAATGATCAAGCTTTCGATGGCAGCGCAACAAAGTCCCGGCTGGGTGCAAGCGCAATC 480
 481 CTTGGTGTTCATGGCTGTTGCAAGGCTGCTGCTGATTCGCGAGGCGCTCCCACTGTTTC 540
 481 CTTGGTGTTCATGGCTGTTGCAAGGCTGCTGCTGATTCGCGAGGCGCTCCCACTGTTTC 540
 541 CGCTACATCGGTGGACCAAAACGACACGCTTCTTCCAGTTTCCAATGATGAACATCATCAAC 600
 541 CGCTACATCGGTGGACCAAAACGACACGCTTCTTCCAGTTTCCAATGATGAACATCATCAAC 600
 601 GGTGGCGCTCAGCTGACTCGGTGTTGAGTTTCAGGAAATTCATGATCGTCCAAATCGGT 660
 601 GGTGGCGCTCAGCTGACTCGGTGTTGAGTTTCAGGAAATTCATGATCGTCCAAATCGGT 660
 661 GCAGAGACCTTCTCTGAGGCTCTCCGCAACGGCGGAGGTCTACACGCACTGAACTCC 720
 661 GCAGAGACCTTCTCTGAGGCTCTCCGCAACGGCGGAGGTCTACACGCACTGAACTCC 720
 721 GTCATCAAGGAAAGGCGCTTCCACCGGACTTGGCGATGAGGCGGCTTCGCTCCTTCC 780
 721 GTCATCAAGGAAAGGCGCTTCCACCGGACTTGGCGATGAGGCGGCTTCGCTCCTTCC 780
 781 GTGGCTTCCACCGTGAAGGCTCTTACCTTATGTTAAGGCAATTCGAGAAAGGCTGCGTTC 840
 781 GTGGCTTCCACCGTGAAGGCTCTTACCTTATGTTAAGGCAATTCGAGAAAGGCTGCGTTC 840
 841 ACCCGAGGCAAGACATCGCTTCTGCTCGAGGTTGCTTCTTCCCTGAGTTCTTCAAGGAC 900
 841 ACCCGAGGCAAGACATCGCTTCTGCTCGAGGTTGCTTCTTCCCTGAGTTCTTCAAGGAC 900
 901 GGCACCTTACCACCTTCGAAGGTGGCCAGCACTCCGCGAGCTGAGATGGCAACGTTTACGCT 960
 901 GGCACCTTACCACCTTCGAAGGTGGCCAGCACTCCGCGAGCTGAGATGGCAACGTTTACGCT 960
 961 GAGTCTGTTGAGCGTTACCAATGTTCTCATTCAGGACCCACTGCGAGGAAGATGACTGG 1020
 961 GAGTCTGTTGAGCGTTACCAATGTTCTCATTCAGGACCCACTGCGAGGAAGATGACTGG 1020
 1021 GAGGTTTACACCAACCTTCAGCGCAACCATCGGCGCAAGGTTTTCAGATCGTTGGCGAGC 1080
 1021 GAGGTTTACACCAACCTTCAGCGCAACCATCGGCGCAAGGTTTTCAGATCGTTGGCGAGC 1080
 1081 TTCTTTCGTCACCAACCTTCAGCGCTTGAAGGAGGCAATCGCTAAGAGGCTTGCACATCC 1140
 1081 TTCTTTCGTCACCAACCTTCAGCGCTTGAAGGAGGCAATCGCTAAGAGGCTTGCACATCC 1140
 1141 ATCTGTTAAGGTGAACAGATCGGTACCTTCACGAGACCTTCGACGCTGTCGACATG 1200
 1141 ATCTGTTAAGGTGAACAGATCGGTACCTTCACGAGACCTTCGACGCTGTCGACATG 1200
 1201 GCTCACCGCGAGGCTACATCTCATGATGTTCCCAACCGTTCCGTTGAGACCGGAGCAC 1260
 1201 GCTCACCGCGAGGCTACATCTCATGATGTTCCCAACCGTTCCGTTGAGACCGGAGCAC 1260

1261 ACCATTGCTGACCTCGCAGTTGACCACTGTGGCAGATCAAGACTGGTGTCTCCAGCA 1320
 1261 ACCATTGCTGACCTCGCAGTTGACCACTGTGGCAGATCAAGACTGGTGTCTCCAGCA 1320
 1321 CGTTCCGACCGTTCGCAAAAGTACAAACAGCTTCTCCGCAATCGAGAGCTGCTTGGCGAC 1380
 1321 CGTTCCGACCGTTCGCAAAAGTACAAACAGCTTCTCCGCAATCGAGAGCTGCTTGGCGAC 1380
 1381 GCCGGCGTCTACCGAGGTTCGCGCAATTCCTCAGGCTTTCAGGCTTAAATAAAGCGCTT 1440
 1381 GCCGGCGTCTACCGAGGTTCGCGCAATTCCTCAGGCTTTCAGGCTTAAATAAAGCGCTT 1440
 1441 TTCGACGCGCGGTAAACCTCAAGGTTCCGGGCGTTCGTTGCTTACTACTGTTACTGTTGT 1500
 1441 TTCGACGCGCGGTAAACCTCAAGGTTTCGGGCGTTCGTTGCTTACTACTGTTACTGTTGT 1500
 1501 GACTATGATCGAGGATTCGCAAGAGCAAGAACTATAAAGGCTTGTTCCTGCTCT 1560
 1501 GACTATGATCGAGGATTCGCAAGAGCAAGAACTATAAAGGCTTGTTCCTGCTCT 1560
 1561 CAAGCAGGGAACGTGCTT 1578
 1561 CAAGCAGGGAACGTGCTT 1578

RESULT 3
 US-09-860-768-1
 ; Sequence 1, Application US/09860768
 ; Patent No. US20020082403A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Mockel, Bettina
 ; APPLICANT: Pfefferle, Walter
 ; APPLICANT: Hermann, Thomas
 ; APPLICANT: Pohler, Alfred
 ; APPLICANT: Kalinowski, Jorn
 ; APPLICANT: Bathe, Brigitte
 ; TITLE OF INVENTION: New Nucleotide Sequences that Code for the Eno Gene
 ; FILE REFERENCE: 21123/278404
 ; CURRENT APPLICATION NUMBER: US/09/860,768
 ; CURRENT FILING DATE: 2001-05-21
 ; NUMBER OF SEQ ID NOS: 6
 ; SOFTWARE: Patentin version 3.0
 ; SEQ ID NO 1
 ; LENGTH: 1578
 ; TYPE: DNA
 ; ORGANISM: Corynebacterium glutamicum
 US-09-860-768-1

Query Match 99.9%; Score 1576.4; DB 9; Length 1578;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1577; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 GGTCTGGGATATGGGTAGTTTTCGCCACTTAATTTCAACTGATTCCTCATCGAAACAAGA 60
 1 GGTCTGGGATATGGGTAGTTTTCGCCACTTAATTTCAACTGATTCCTCATCGAAACAAGA 60
 61 TTCTGTCACCAATTTGGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 120
 61 TTCTGTCACCAATTTGGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 120
 121 TTAGCTCCCAAGTTGGCATAGGAGCCACAGTGGCTGAAATCATGACGTTATTCGCTCGC 180
 121 TTAGCTCCCAAGTTGGCATAGGAGCCACAGTGGCTGAAATCATGACGTTATTCGCTCGC 180
 181 GAAATTTCTGACTCCCGCGGTAAACCAACCGTTCGAGGAGAGGTTTTCCTGGATGACGGT 240
 181 GAAATTTCTGACTCCCGCGGTAAACCAACCGTTCGAGGAGAGGTTTTCCTGGATGACGGT 240
 241 TCCACCGGTTCGAGGTTCCTCATCGGCGCATCCACCGGCTCCACCGGCTCATGAG 300
 241 TCCACCGGTTCGAGGTTCCTCATCGGCGCATCCACCGGCTCCACCGGCTCATGAG 300
 301 CTGCGTGAACGTTGGCGATCGCTACCTGGGCAAGGGCGTTTTTGAAGGCGATTGAAGAGTC 360

```
Db 301 |||||CTGGTGACGGTGGCGATCGTACTCTGGCGAGGGCGTTTTGAAGGCAGTTGAAAAAGTGC 360
Qy 361 AACGAAGAAATCGCGCAGCGACTCGCTGGCCTAGAGGCTGACGATCAGCGCCTCATCGAC 420
Db 361 AACGAAGAAATCGCGCAGCGAGCTCGCTGGCCTAGAGGCTGAGGATCAGCGCCTCATCGAC 420
Qy 421 GAAGCAATGATCAAGCTTGATGGCAGCCGCCAACAAGTCCCGCCTGGGTGGAACGCAATC 480
Db 421 GAAGCAATGATCAAGCTTGATGGCAGCCGCCAACAAGTCCCGCCTGGGTGGAACGCAATC 480
Qy 481 CTTGGTGTTCATCGGCTGTGTGAAAAGCGTCTGCTGATTCGCGAGGCGCTCCCACTGTC 540
Db 481 CTTGGTGTTCATCGGCTGTGTGAAAAGCGTCTGCTGATTCGCGAGGCGCTCCCACTGTC 540
Qy 541 CGCTACATCGGTGGACCAAAAGCGACACGTTTCTCCAGTTCCAATGATGAACATCATCAAC 600
Db 541 CGCTACATCGGTGGACCAAAAGCGACACGTTTCTCCAGTTCCAATGATGAACATCATCAAC 600
Qy 601 GGTGGCGCTCAGCTGACTCGGTGTTGACGTTCAGGAATTCATGATCGCTCCATCGGT 660
Db 601 GGTGGCGCTCAGCTGACTCGGTGTTGACGTTTCCAGGAATTCATGATCGCTCCATCGGT 660
Qy 661 GCAGAGACCTTCTCTGAGGCTCTCCGCAACGGCGCGAGGCTCTACCAACGCACTGAAGTCC 720
Db 661 GCAGAGACCTTCTCTGAGGCTCTCCGCAACGGCGCGAGGCTCTACCAACGCACTGAAGTCC 720
Qy 721 GTCATCAAGAAAAGGCCCTGTCCACCGACTTGCGCGATGAGGGCGCTTCGCTCCCTCC 780
Db 721 GTCATCAAGAAAAGGCCCTGTCCACCGACTTGCGCGATGAGGGCGCTTCGCTCCCTCC 780
Qy 781 GTGGCTCCACCGCTGAGGCTCTTGACCTTATCGTTAAGCAATCGAAGAGGCTGGCTTC 840
Db 781 GTGGCTCCACCGCTGAGGCTCTTGACCTTATCGTTAAGCAATCGAAGAGGCTGGCTTC 840
Qy 841 ACCCAGGCAAGGACATCGCTTTGCTCTGGAAGTTCCTTCTCTGAGTTCTTCAAGGAC 900
Db 841 ACCCAGGCAAGGACATCGCTTTGCTCTGGAAGTTCCTTCTCTGAGTTCTTCAAGGAC 900
Qy 901 GGCACCTTACCACCTCGAAGTGGCCAGCAGCTCCGACGCTGAGATGGCAAAAGTTAGCT 960
Db 901 GGCACCTTACCACCTCGAAGTGGCCAGCAGCTCCGACGCTGAGATGGCAAAAGTTAGCT 960
Qy 961 GAGTCGTTGACGCGTACCCAACTCGTCTCCATCGAGACCCACTGCAAGGAAGATGACTGG 1020
Db 961 GAGTCGTTGACGCGTACCCAACTCGTCTCCATCGAGGACCCACTGCAAGGAAGATGACTGG 1020
Qy 1021 GAGGTTTACACCACTTCAACCGCAACCATCGGCAACAAGTTTCAAGTTCGTTGGCGACGAC 1080
Db 1021 GAGGTTTACACCACTTCAACCGCAACCATCGGCAACAAGTTTCAAGTTCGTTGGCGACGAC 1080
Qy 1081 TTCTTGTTCACCAACCTCGAGCGCTGAAGGAGGATCGCTAAGAAGGCTGCAACTCC 1140
Db 1081 TTCTTGTTCACCAACCTCGAGCGCTGAAGGAGGATCGCTAAGAAGGCTGCAACTCC 1140
Qy 1141 ATCTTGGTTAAGTGAACCAAGATCGGTACCTTCAACGAGACCTTCAACGCTGTCGACATG 1200
Db 1141 ATCTTGGTTAAGTGAACCAAGATCGGTACCTTCAACGAGACCTTCAACGCTGTCGACATG 1200
Qy 1201 GCTCAGCGCGAGGCTACACTCGATGATGTCCACCGTTCCGGTGAGACCGAGGACACC 1260
Db 1201 GCTCAGCGCGAGGCTACACTCGATGATGTCCACCGTTCCGGTGAGACCGAGGACACC 1260
Qy 1261 ACCATTGCTGACCTCGCAGTTGCACTCAACTGTGGCCAGATCAAGACTGGTGTCCAGCA 1320
Db 1261 ACCATTGCTGACCTCGCAGTTGCACTCAACTGTGGCCAGATCAAGACTGGTGTCCAGCA 1320
Qy 1321 CGTTCGACCGGTGCGAAAGTACAACAGCTTCTCCGATCGAGCAGCTGCTTGGCGAC 1380
Db 1321 CGTTCGACCGGTGCGAAAGTACAACAGCTTCTCCGATCGAGCAGCTGCTTGGCGAC 1380
Qy 1381 GCCGGCGTCTACCGAGGTGCGAGGATTCACCGCTTTCAGGGCTTAATTAAGGCGCTT 1440
Db 1381 GCCGGCGTCTACCGAGGTGCGAGGATTCACCGCTTTCAGGGCTTAATTAAGGCGCTT 1440
```

```
Db 1381 GCCGGCGTCTACCGAGGTGCGAGGATTCACCGCTTTCAGGGCTTAATTAAGGCGCTT 1440
Qy 1441 TTCCAGCCCGCGTAAACCTCAAGGTTCCCGGGCTCGTTGCTTACTACTGTTACTGGTGT 1500
Db 1441 TTCCAGCCCGCGTAAACCTCAAGGTTCCCGGGCTCGTTGCTTACTACTGTTACTGGTGT 1500
Qy 1501 GACTATGATCGAGGATTTATGGCAAGCAGAGAAACTATAAAGGCTTTGCTCTGCT 1560
Db 1501 GACTATGATCGAGGATTTATGGCAAGCAGAGAAACTATAAAGGCTTTGCTCTGCT 1560
Qy 1561 CAAGCAGGGAACGTGCTT 1578
Db 1561 CAAGCAGGGAACGTGCTT 1578

RESULT 4
US-10-728-947-1
; Sequence 1, Application US/10728947
; Publication No. US20040220394A1
; GENERAL INFORMATION:
; APPLICANT: Mockel, Bettina
; APPLICANT: Pfefferle, Walter
; APPLICANT: Hermann, Thomas
; APPLICANT: Pohler, Alfred
; APPLICANT: Kalinowski, Jorn
; APPLICANT: Bathe, Brigitte
; TITLE OF INVENTION: New Nucleotide Sequences that Code for the Eno Gene
; FILE REFERENCE: 21123/278404
; CURRENT APPLICATION NUMBER: US/10/728,947
; PRIOR FILING DATE: 2003-12-08
; PRIOR APPLICATION NUMBER: US/09/860,768
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 1578
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
US-10-728-947-1

Query Match 99.9%; Score 1576.4; DB 20; Length 1578;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1577; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCGTGGGATATGGGTAGTTTTCGCCACTTAATTTCAACTGATTCGCTCATCGAAAACAAG 60
Db 1 GCGTGGGATATGGGTAGTTTTCGCCACTTAATTTCAACTGATTCGCTCATCGAAAACAAG 60
Qy 61 TTCGTGCAACAATTTGGGTGTAGACGTGATTGAAGACATTTGATCAGTGAATAATTTCTAG 120
Db 61 TTCGTGCAACAATTTGGGTGTAGACGTGATTGAAGACATTTGATCAGTGAATAATTTCTAG 120
Qy 121 TTAGTCTCCAAAGTTGGCATAGGAGGCCACAGTGGCTGAAATCATGCAGTATTCGCTCGC 180
Db 121 TTAGTCTCCAAAGTTGGCATAGGAGGCCACAGTGGCTGAAATCATGCAGTATTCGCTCGC 180
Qy 181 GAAATTTCTCGACTCCCGCGTAAACCAACCGTCGAGGAGAGGTTTCTCGATGACGGT 240
Db 181 GAAATTTCTCGACTCCCGCGTAAACCAACCGTCGAGGAGAGGTTTCTCGATGACGGT 240
Qy 241 TCCCAGGTTGTCGAGGTTTCCATCCGGCGCATCCACCGGCTCCACGAGGCTCATGAG 300
Db 241 TCCCAGGTTGTCGAGGTTTCCATCCGGCGCATCCACCGGCTCCACGAGGCTCATGAG 300
Qy 301 CTGCGTGACGGTGGCGATCGCTACTTGGGCAAGGGGCTTTGAAAGCAGTTGAAAACGTC 360
Db 301 CTGCGTGACGGTGGCGATCGCTACTTGGGCAAGGGGCTTTGAAAGCAGTTGAAAACGTC 360
Qy 361 AACGAAGAAATCGCGCAGCGACTCGCTGGCCTAGAGGCTGACGATCAGCGCCTCATCGAC 420
Db 361 AACGAAGAAATCGCGCAGCGACTCGCTGGCCTAGAGGCTGACGATCAGCGCCTCATCGAC 420
Qy 421 GAAGCAATGATCAAGCTTGATGGCAGCCGCCAACAAGTCCCGCCTGGGTGGAACGCAATC 480
Db 421 GAAGCAATGATCAAGCTTGATGGCAGCCGCCAACAAGTCCCGCCTGGGTGGAACGCAATC 480
Qy 481 CTTGGTGTTCATCGGCTGTGTGAAAAGCGTCTGCTGATTCGCGAGGCGCTCCCACTGTC 540
Db 481 CTTGGTGTTCATCGGCTGTGTGAAAAGCGTCTGCTGATTCGCGAGGCGCTCCCACTGTC 540
```

Db 421 GAAGCAATGATCAAGCTTGTATGGCACCGCCAAACAAGTCCGGCTGGGTGCAAAACGCAATC 480
QY 481 CTTGGGTGTTTCCATGGCTGTTGCAAAAGCGTCTGCTGATTTCCGAGGCGCTCCACACTGTTTC 540
Db 481 CTTGGGTGTTTCCATGGCTGTTGCAAAAGCGTCTGCTGATTTCCGAGGCGCTCCACACTGTTTC 540
QY 541 CGCTACATCGGTGGACCAACGCAACGACGTTCTTCCAGTTTCCAATGATGAACATCATCAAC 600
Db 541 CGCTACATCGGTGGACCAACGCAACGACGTTCTTCCAGTTTCCAATGATGAACATCATCAAC 600
QY 601 GGTGGCGCTCACGCTGACTCCGTTGTTGACGTTTCAGGAATTCATGATCGCTCCAATCGGT 660
Db 601 GGTGGCGCTCACGCTGACTCCGTTGTTGACGTTTCAGGAATTCATGATCGCTCCAATCGGT 660
QY 661 GCAGAGACCTTCTCTGAGGCTCTCCGCAACGCGCGGAGGTCTACCAACGCACTGMAATGCC 720
Db 661 GCAGAGACCTTCTCTGAGGCTCTCCGCAACGCGCGGAGGTCTACCAACGCACTGMAATGCC 720
QY 721 GTCATCAAGGAAAGGCGCTGTCCACGAGCTTTGGCGATGAGGCGGCTTCGCTCCTTCC 780
Db 721 GTCATCAAGGAAAGGCGCTGTCCACGAGCTTTGGCGATGAGGCGGCTTCGCTCCTTCC 780
QY 781 GTCGGCTCCACCGTGAGGCTCTTGACCTTATCGTTAAGGCAATCGAAGAGGCTGGCTTC 840
Db 781 GTCGGCTCCACCGTGAGGCTCTTGACCTTATCGTTAAGGCAATCGAAGAGGCTGGCTTC 840
QY 841 ACCCGAGGCAAGACATCGCTCTGCTCTGAGAGTTCCTCTCTGAGTTCCTCAAGGAC 900
Db 841 ACCCGAGGCAAGACATCGCTCTGCTCTGAGAGTTCCTCTCTGAGTTCCTCAAGGAC 900
QY 901 GGCACCTTACCACCTCGAAGGTGGCGAGCACTCCGCACTGAGATGGCAACGTTTACGCT 960
Db 901 GGCACCTTACCACCTCGAAGGTGGCGAGCACTCCGCACTGAGATGGCAACGTTTACGCT 960
QY 961 GAGCTGTTGACCGGTACCAATCGTCTCATCGAGACCCACTGCGAGGAAGATGACTGG 1020
Db 961 GAGCTGTTGACCGGTACCAATCGTCTCATCGAGACCCACTGCGAGGAAGATGACTGG 1020
QY 1021 GAGGTTTACACCAACCTCACCGCAACCATCGGCGACAAGGTTTCAGATCGTTGGCGAGC 1080
Db 1021 GAGGTTTACACCAACCTCACCGCAACCATCGGCGACAAGGTTTCAGATCGTTGGCGAGC 1080
QY 1081 TTCTTCGTCAACCAACCTCGAGCGCTGAAGGAGGCACTCGTAAGAGGCTGCAACTCC 1140
Db 1081 TTCTTCGTCAACCAACCTCGAGCGCTGAAGGAGGCACTCGTAAGAGGCTGCAACTCC 1140
QY 1141 ATCCTGGTTAAGGTGAACCAAGATCGGTACCCCTCACCGAGACCTTCGACGCTGTCGACATG 1200
Db 1141 ATCCTGGTTAAGGTGAACCAAGATCGGTACCCCTCACCGAGACCTTCGACGCTGTCGACATG 1200
QY 1201 GCTACCGCGCAGCTACACCTCCATGATGTCCACCGTTCCGGTGAGACGAGGACACC 1260
Db 1201 GCTACCGCGCAGCTACACCTCCATGATGTCCACCGTTCCGGTGAGACGAGGACACC 1260
QY 1261 ACCATTGCTGACCTCGCAGTTGCACTCAACTGTGGCGAGATCAAGACTGGTGTCCAGCA 1320
Db 1261 ACCATTGCTGACCTCGCAGTTGCACTCAACTGTGGCGAGATCAAGACTGGTGTCCAGCA 1320
QY 1321 CGTTCCGACCGGTGCGCAAGTACAAACGAGTTCCTCGCATCGAGCAGCTGCTTGGCGAC 1380
Db 1321 CGTTCCGACCGGTGCGCAAGTACAAACGAGTTCCTCGCATCGAGCAGCTGCTTGGCGAC 1380
QY 1381 GCGGCGCTTACCGAGGTCCAGGCAATTCACGCTTTTCAGGGCTAAATAAAGCGCTT 1440
Db 1381 GCGGCGCTTACCGAGGTCCAGGCAATTCACGCTTTTCAGGGCTAAATAAAGCGCTT 1440
QY 1441 TTCGACGCGCGGTAACTCAAGGTTCGCGGCGCTCGTTGCTTACTACTGTTACTGCTGT 1500
Db 1441 TTCGACGCGCGGTAACTCAAGGTTCGCGGCGCTCGTTGCTTACTACTGTTACTGCTGT 1500
QY 1501 GACTATGATCGAGGATTTATGGCAAGCAGAAAGAACTCATAAAGGCGCTTGTTCCTGTCT 1560

Db 1501 GACTATGATCGAGGATTTATGGCAAGCAGAAAGAACTCATAAAGGCGCTTGTTCCTGTCT 1560
QY 1561 CAAGCAGGGAACGTGCTT 1578
Db 1561 CAAGCAGGGAACGTGCTT 1578
RESULT 5
US-09-738-626-1
; Sequence 1, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738, 626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 1
; LENGTH: 3309400
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-1

Query Match 99.9%; Score 1576.4; DB 9; Length 3309400;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1577; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GCGTGGGATATGGGTAGTCTTCGCCACTAAATTTCAACTGATTTGCCCTCATCGAACAAGA 60
Db 1034799 GCGTGGGATATGGGTAGTCTTCGCCACTAAATTTCAACTGATTTGCCCTCATCGAACAAGA 1034858
QY 61 TTCGTGCAACAATTTGGGTGTAGACGTGATTGAAGACATTTTGATCACGTGAATATTTCTAG 120
Db 1034859 TTCGTGCAACAATTTGGGTGTAGACGTGATTGAAGACATTTTGATCACGTGAATATTTCTAG 1034918
QY 121 TTAGCTCCCAAGTTGGCATAGGAGGCCACAGTGGCTGAAATCATGACGTATTCGCTCGC 180
Db 1034919 TTAGCTCCCAAGTTGGCATAGGAGGCCACAGTGGCTGAAATCATGACGTATTCGCTCGC 1034978
QY 181 GAAATTCCTGACTCCCGCGGTAAACCAACCGTCCAGGACAGGCTTTTCCTGGATGACGT 240
Db 1034979 GAAATTCCTGACTCCCGCGGTAAACCAACCGTCCAGGACAGGCTTTTCCTGGATGACGT 1035038
QY 241 TCCACCGGTGTCCAGGTGTTCCATCCGGCGCATCCACCGCGCTCCACGAGGCTCATGAG 300
Db 1035039 TCCACCGGTGTCCAGGTGTTCCATCCGGCGCATCCACCGCGCTCCACGAGGCTCATGAG 1035098
QY 301 CTGCGTGA CGGTGGCGATCGCTACCTTGGGCAAGGCGTCTTTGAAGGAGTGTGAAACGTC 360
Db 1035099 CTGCGTGA CGGTGGCGATCGCTACCTTGGGCAAGGCGTCTTTGAAGGAGTGTGAAACGTC 1035158
QY 361 AACGAGAATCGGCGACGAGCTCGCTGGCTAGAGGCTGACCATCAGCGCTCATCGAC 420
Db 1035159 AACGAGAATCGGCGACGAGCTCGCTGGCTAGAGGCTGACCATCAGCGCTCATCGAC 1035218
QY 421 GAAGCAATGATCAAGCTTGTATGGCACCGGCCAAAGTCCCGCTGGGTGCAAAACGCAATC 480

Db	1035219	GAAGCAATGATCAAGCTTTGATGGACCGCCAAACAGTCCCGCTGGGTGCAAAACGCAATC	1035278
Qy	481	CTTGGTGTTCCTATGGCTGTGTCGAAAGGCTGTGCTGATTCGAGGCTCCCACTGTTC	540
Db	1035279	CTTGGTGTTCCTATGGCTGTGTCGAAAGGCTGTGCTGATTCGAGGCTCCCACTGTTC	1035338
Qy	541	CGCTACATCGGTGACCAACAGCACAGTTCCTCCAGTTCCAGTTCCCAATGATGAACATCATCAAC	600
Db	1035339	CGCTACATCGGTGACCAACAGCACAGTTCCTCCAGTTCCAGTTCCCAATGATGAACATCATCAAC	1035398
Qy	601	GGTGGCGCTCACGCTGACTCCGCTGTGAGCTTCAGGAAATTCATGATCGCTCCAATCGGT	660
Db	1035399	GGTGGCGCTCACGCTGACTCCGCTGTGAGCTTCAGGAAATTCATGATCGCTCCAATCGGT	1035458
Qy	661	GCAGAGACCTTCTCTGAGGCTCTCCGCAAGCGGCGAGGTGTACACGCACTGAAGTCC	720
Db	1035459	GCAGAGACCTTCTCTGAGGCTCTCCGCAAGCGGCGAGGTGTACACGCACTGAAGTCC	1035518
Qy	721	GTCAATCAAGAAAAGGCTGTCCACCGACTTGGCGATCAGGCGGCTTCGCTCCTTCC	780
Db	1035519	GTCAATCAAGAAAAGGCTGTCCACCGACTTGGCGATCAGGCGGCTTCGCTCCTTCC	1035578
Qy	781	GTGGCTTCCACCGCTGAGGCTCTTGACCTTATCGTTAAGCAATCGAAGAGGCTGGCTTC	840
Db	1035579	GTGGCTTCCACCGCTGAGGCTCTTGACCTTATCGTTAAGCAATCGAAGAGGCTGGCTTC	1035638
Qy	841	ACCCAGGCAAGACATCGCTCTTGTCTCGAGCTTGGCGATCAGGCGGCTTCGCTCCTTCC	900
Db	1035639	ACCCAGGCAAGACATCGCTCTTGTCTCGAGCTTGGCGATCAGGCGGCTTCGCTCCTTCC	1035698
Qy	901	GGCACTTACCACCTTCGAAGGTGGCGACACTCCGCACTGAGATGGCAACGTTTACGCT	960
Db	1035699	GGCACTTACCACCTTCGAAGGTGGCGACACTCCGCACTGAGATGGCAACGTTTACGCT	1035758
Qy	961	GAGCTCGTTGACGGTACCCTTGTCTTCATCGAGGACCCACTGTCAGGAAGTGAAGTGG	1020
Db	1035759	GAGCTCGTTGACGGTACCCTTGTCTTCATCGAGGACCCACTGTCAGGAAGTGAAGTGG	1035818
Qy	1021	GAGGTTTACACCAACTTCCCGCAACCATCGGCGACAGGTTTCCAGTTCGTCGGGACGAC	1080
Db	1035819	GAGGTTTACACCAACTTCCCGCAACCATCGGCGACAGGTTTCCAGTTCGTCGGGACGAC	1035878
Qy	1081	TTCTTCGTACCAACCTCGAGCGCTGAAGAGGGCATCGCTAAGAGGCTGCAACTCC	1140
Db	1035879	TTCTTCGTACCAACCTCGAGCGCTGAAGAGGGCATCGCTAAGAGGCTGCAACTCC	1035938
Qy	1141	ATCCTGGTTAAGGTGAACCAAGATCGGTACCTCACCGAGACCTTCGACGCTGTGCAATG	1200
Db	1035939	ATCCTGGTTAAGGTGAACCAAGATCGGTACCTCACCGAGACCTTCGACGCTGTGCAATG	1035998
Qy	1201	GCTCACGGCGAGGCTACACCTCCATGATGTCCACCGTTCGGTCCGACCGAGGACAC	1260
Db	1035999	GCTCACGGCGAGGCTACACCTCCATGATGTCCACCGTTCGGTCCGACCGAGGACAC	1036058
Qy	1261	ACCAATGCTGACCTCGCAGTTGCACTCAACTGTGGCCAGATCAAGACTGTGTCTCCAGCA	1320
Db	1036059	ACCAATGCTGACCTCGCAGTTGCACTCAACTGTGGCCAGATCAAGACTGTGTCTCCAGCA	1036118
Qy	1321	CGTTCCGACCGTGTCCGAAAGTACACCACTTCTCCGATCGAGGAGCTGTGTGGCGAC	1380
Db	1036119	CGTTCCGACCGTGTCCGAAAGTACACCACTTCTCCGATCGAGGAGCTGTGTGGCGAC	1036178
Qy	1381	GC CGGCGTCTACGAGGTCCGACGATTCCTCCAGCTTTTCAGGGCTAAATAAAGCGCTT	1440
Db	1036179	GC CGGCGTCTACGAGGTCCGACGATTCCTCCAGCTTTTCAGGGCTAAATAAAGCGCTT	1036238
Qy	1441	TTCCGACCGCGGTAACTCAAGGTTCCGCGGCTCGTTGCTTACTACTGTTACTCGGTGT	1500
Db	1036239	TTCCGACCGCGGTAACTCAAGGTTCCGCGGCTCGTTGCTTACTACTGTTACTCGGTGT	1036298
Qy	1501	GACTATGATCGAGGATTTATGGCAAGACAGAGAAACTATAAAGCGCTTGTCTCTGTCT	1560

Db	1036299	GACTATGATCGAGGATTTATGGCAAGACAGAGAAACTATAAAGCGCTTGTCTCTGTCT	1036358
Qy	1561	CAAGCAGGGAACGTGCTT 1578	
Db	1036359	CAAGCAGGGAACGTGCTT 1036376	
RESULT 6			
US-10-494-836-7			
; Sequence 7, Application US/10494836			
; Publication No. US20050014233A1			
; GENERAL INFORMATION:			
; APPLICANT: Zelder, Oskar			
; APPLICANT: Pompejus, Markus			
; APPLICANT: Schroder, Hartwig			
; APPLICANT: Kroger, Burkhard			
; APPLICANT: Klopffrogge, Corinna			
; APPLICANT: Haberhauer, Gregor			
; TITLE OF INVENTION: Genes coding for proteins of carbon metabolism and energy product			
; FILE REFERENCE: BGI-167US			
; CURRENT APPLICATION NUMBER: US/10/494,836			
; CURRENT FILING DATE: 2004-05-05			
; PRIOR APPLICATION NUMBER: PCT/EP02/12135			
; PRIOR FILING DATE: 2002-10-31			
; PRIOR APPLICATION NUMBER: DE 101 54 270.4			
; PRIOR FILING DATE: 2001-11-05			
; NUMBER OF SEQ ID NOS: 116			
; SEQ ID NO 7			
; LENGTH: 1405			
; TYPE: DNA			
; ORGANISM: Corynebacterium glutamicum			
; FEATURE:			
; NAME/KEY: CDS			
; LOCATION: (101)..(1375)			
; OTHER INFORMATION: RXA00235			
US-10-494-836-7			

Query Match		88.8%	Score 1401.8;	DB 21;	Length 1405;
Best Local Similarity		99.9%	Pred. No. 0;		
Matches 1403;		Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;
Qy	51	CGAAACAAGATTCTGGTGTAGACAAATTTGGGTGTAGACAAATTTGAAGACATTTGATCAGTGA	110		
Db	1	CGAAACAAGATTCTGGTGTAGACAAATTTGGGTGTAGACAAATTTGAAGACATTTGATCAGTGA	60		
Qy	111	ATAATTCTAGTTAGCTCCCAAGTTGGCATAGAGGCCACAGTGGCTGAAATCATGCACGT	170		
Db	61	ATAATTCTAGTTAGCTCCCAAGTTGGCATAGAGGCCACAGTGGCTGAAATCATGCACGT	120		
Qy	171	ATTTCGCTCGCAAAATTTCTCGACTCCCGCGTAAACCAACCGTCCGAGGACAGAGTTTTCCT	230		
Db	121	ATTTCGCTCGCAAAATTTCTCGACTCCCGCGTAAACCAACCGTCCGAGGACAGAGTTTTCCT	180		
Qy	231	GGATGACGGTTCCTCCAGTGTGCGAGTGTTCATTCGCGGCATCCACCGGGTCCACGA	290		
Db	181	GGATGACGGTTCCTCCAGTGTGCGAGTGTTCATTCGCGGCATCCACCGGGTCCACGA	240		
Qy	291	GGCTCATGAGCTCCGTCGAGCGTGGCGATCGCTACCTGGGCAAGGGCGTTTGAAGGCAGT	350		
Db	241	GGCTCATGAGCTCCGTCGAGCGTGGCGATCGCTACCTGGGCAAGGGCGTTTGAAGGCAGT	300		
Qy	351	TGAAACAAGTTCACGAAGAAATCGGCGACGAGCTCGCTGGCCTAGAGGCTGACGATCAGCG	410		
Db	301	TGAAACAAGTTCACGAAGAAATCGGCGACGAGCTCGCTGGCCTAGAGGCTGACGATCAGCG	360		
Qy	411	CCTCATCGAAGCAATGATCAAGTGTGATGGCACCGCCAAACAAGTCCCGCTGGGTGC	470		
Db	361	CCTCATCGAAGCAATGATCAAGTGTGATGGCACCGCCAAACAAGTCCCGCTGGGTGC	420		
Qy	471	AAACCCAACTCCTGGTGTTCATCGCTGTTTCCAAAGGCTGCTGATTCGACAGGCTT	530		
Db	421	AAACCCAACTCCTGGTGTTCATCGCTGTTTCCAAAGGCTGCTGATTCGACAGGCTT	480		

QY 531 CCCACTGTTCCGCTACATCGGTGGACCAACGACACAGTTCCTCCAGTTCCTCAATGATGAA 590
 Db 481 CCCACTGTTCCGCTACATCGGTGGACCAACGACACAGTTCCTCCAGTTCCTCAATGATGAA 540
 QY 591 CATCATCAAGGTTGGGCTCACGCTGACTCCGGTGTGAGTTCAGGAATTCATGATCGC 650
 Db 541 CATCATCAAGGTTGGGCTCACGCTGACTCCGGTGTGAGTTCAGGAATTCATGATCGC 600
 QY 651 TCCAAATCGGTGCAGAGACCTTCTCTGAGGCTCTCCGCAACGGCGGGAGGTCTACCAACGC 710
 Db 601 TCCAAATCGGTGCAGAGACCTTCTCTGAGGCTCTCCGCAACGGCGGGAGGTCTACCAACGC 660
 QY 711 ACTGAAGTCCGTTCATCAAGAAAGGCGCTGTCCACCGGACTTTGGCGATGAGGCGCGCTT 770
 Db 661 ACTGAAGTCCGTTCATCAAGAAAGGCGCTGTCCACCGGACTTTGGCGATGAGGCGCGCTT 720
 QY 771 CGCTCTCTCGGTGGCTTCCACCGGTGAGGCTCTTGACCTTATCGTTAAGGCAATCGAGAA 830
 Db 721 CGCTCTCTCGGTGGCTTCCACCGGTGAGGCTCTTGACCTTATCGTTAAGGCAATCGAGAA 780
 QY 831 GGCTGGCTTCAACCCAGGCAAGGACATCGCTCTTGCTCTCGAGTTCGTTCTCTCTGAGTT 890
 Db 781 GGCTGGCTTCAACCCAGGCAAGGACATCGCTCTTGCTCTCGAGTTCGTTCTCTCTGAGTT 840
 QY 891 CTTTCAAGGACGGCACTTACCACTTCGAAGGTGGCCAGCACTCCGCGAGCTGAGATGGCAAA 950
 Db 841 CTTTCAAGGACGGCACTTACCACTTCGAAGGTGGCCAGCACTCCGCGAGCTGAGATGGCAAA 900
 QY 951 CGTTTACGCTGAGCTGTTGACGGGTACCAATCGTCTCCATCGAGTTCGATGAGGA 1010
 Db 901 CGTTTACGCTGAGCTGTTGACGGGTACCAATCGTCTCCATCGAGTTCGATGAGGA 960
 QY 1011 AGATGACTGGAGGGTTTACACCAACCTTACCGCAACCATCGGCGCAAGGTTCCAGATCGT 1070
 Db 961 AGATGACTGGAGGGTTTACACCAACCTTACCGCAACCATCGGCGCAAGGTTCCAGATCGT 1020
 QY 1071 TGGCGAGCACTTCTCGTCAACCAACCTTGAGCGCTTGAAGGAGGATCGCTAAGAAGGC 1130
 Db 1021 TGGCGAGCACTTCTCGTCAACCAACCTTGAGCGCTTGAAGGAGGATCGCTAAGAAGGC 1080
 QY 1131 TGGCAACTCGATCTGTTAAGGTGAACAGATCGGTACCTTCCAGAGACCTTCCAGCGC 1190
 Db 1081 TGGCAACTCGATCTGTTAAGGTGAACAGATCGGTACCTTCCAGAGACCTTCCAGCGC 1140
 QY 1191 TGTGCACTGAGCTCACCGCGAGGCTACACTCCATGATGTCCTCCAGCTTCGCGTGAGAC 1250
 Db 1141 TGTGCACTGAGCTCACCGCGAGGCTACACTCCATGATGTCCTCCAGCTTCGCGTGAGAC 1200
 QY 1251 CGAGGACACCACTTGTGACCTCGCAGTTCGACTCAACTGTGGCCAGATCAAGACTGG 1310
 Db 1201 CGAGGACACCACTTGTGACCTCGCAGTTCGACTCAACTGTGGCCAGATCAAGACTGG 1260
 QY 1311 TGTCTCAGCACTGTCGACCGGTGCGAAGTGAACAGCTTCTCCGATTCGAGAGCT 1370
 Db 1261 TGTCTCAGCACTGTCGACCGGTGCGAAGTGAACAGCTTCTCCGATTCGAGAGCT 1320
 QY 1371 GCTTGGCGAGCGCGGCTTACCGAGGTTCGAGGATTCGAGCTTCCAGGCTTTCAGGGCTAAAT 1430
 Db 1321 GCTTGGCGAGCGCGGCTTACCGAGGTTCGAGGATTCGAGCTTCCAGGCTTTCAGGGCTAAAT 1380
 QY 1431 AAAAGCGCTTTCGAGCGCGGCTAA 1455
 Db 1381 AAAAGCGCTTTCGAGCGCGGCTAA 1405

RESULT 7

US-10-781-014-71
 ; Sequence 71, Application US/10781014
 ; Publication No. US20040180408A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Pompejus, Markus
 ; APPLICANT: Kroger, Burkhard
 ; APPLICANT: Schroder, Hartwig

; APPLICANT: Zelder, Oskar
 ; APPLICANT: Haberhauer, Gregor
 ; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
 ; TITLE OF INVENTION: INVOLVED IN CARBON METABOLISM AND ENERGY
 ; TITLE OF INVENTION: PRODUCTION
 ; FILE REFERENCE: BGI-126CPCN
 ; CURRENT FILING DATE: 2004-02-17
 ; PRIOR APPLICATION NUMBER: US/10/781,014
 ; PRIOR FILING DATE: 2000-06-23
 ; PRIOR APPLICATION NUMBER: 60/141,031
 ; PRIOR FILING DATE: 1999-06-25
 ; PRIOR APPLICATION NUMBER: 60/143,208
 ; PRIOR FILING DATE: 1999-07-09
 ; PRIOR APPLICATION NUMBER: 60/151,572
 ; PRIOR FILING DATE: 1999-08-31
 ; PRIOR APPLICATION NUMBER: DE 19931412.8
 ; PRIOR FILING DATE: 1999-07-08
 ; PRIOR APPLICATION NUMBER: DE 19931413.6
 ; PRIOR FILING DATE: 1999-07-08
 ; PRIOR APPLICATION NUMBER: DE 19931419.5
 ; PRIOR FILING DATE: 1999-07-08
 ; PRIOR APPLICATION NUMBER: DE 19931420.9
 ; PRIOR FILING DATE: 1999-07-08
 ; PRIOR APPLICATION NUMBER: DE 19931424.1
 ; PRIOR FILING DATE: 1999-07-08
 ; PRIOR APPLICATION NUMBER: DE 19931428.4
 ; PRIOR FILING DATE: 1999-07-08
 ; Remaining prior application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 784
 ; SEQ ID NO 71
 ; LENGTH: 1398
 ; TYPE: DNA
 ; ORGANISM: Corynebacterium glutamicum
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (101)..(1375)
 ; OTHER INFORMATION: RXA00235
 ; US-10-781-014-71

Query Match 88.4%; Score 1394.8; DB 19; Length 1398;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1396; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 51 CGAAACAAGATTCGTCAACAATTTGGTGTAGACGTGATTGAAGACATTTGATCACGTGA 110
 Db 1 CGAAACAAGATTCGTCAACAATTTGGTGTAGACGTGATTGAAGACATTTGATCACGTGA 60

QY 111 ATAAATCTAGTTAGCTCCCAAGTTGGCATAGGAGGCCACAGTGGCTGAAATCATGCACGT 170
 Db 61 ATAAATCTAGTTAGCTCCCAAGTTGGCATAGGAGGCCACAGTGGCTGAAATCATGCACGT 120

QY 171 ATTGCTCGCGAAATTTCTGACTCCCGCGGTAAACCAACCGTTCGAGGACAGGTTTTCCT 230
 Db 121 ATTGCTCGCGAAATTTCTGACTCCCGCGGTAAACCAACCGTTCGAGGACAGGTTTTCCT 180

QY 231 GGATGAGCGTTCACCGGTGTCCAGTGTTCATCGGCGCATCCACCGGCTCCACGA 290
 Db 181 GGATGAGCGTTCACCGGTGTTCATCGGCGCATCCACCGGCTCCACGA 240

QY 291 GGCTCATGAGTTCGCTGACGCTGGCGATCGCTACCTGGGCAAGGGGCTTTTGAAGGAGT 350
 Db 241 GGCTCATGAGTTCGCTGACGCTGGCGATCGCTACCTGGGCAAGGGGCTTTTGAAGGAGT 300

QY 351 TGAACAACGTCAACGAAGAAATCGGCGACGAGCTCGCTGGCTTAGAGGCTGACGATCAGCG 410
 Db 301 TGAACAACGTCAACGAAGAAATCGGCGACGAGCTCGCTGGCTTAGAGGCTGACGATCAGCG 360

QY 411 CCTCATCGAGGAGCAATGATCAAGCTTATGCGCACCGGCAACAAAGTCCCGCTGGGTGC 470
 Db 361 CCTCATCGAGGAGCAATGATCAAGCTTATGCGCACCGGCAACAAAGTCCCGCTGGGTGC 420

QY 471 AAACGCAATCTTTGGTGTTCATGCTGTTGCAAGGCTGCTGCTGATTCCCGCAGSCCT 530

```

Db 421 AAAGCAATCTTGGTGTTCATGGCTGTGCAAGGCTGCTGCTGATTCGACGGCT 480
QY 531 CCACATGTTCCGGTACATCCGGTGGAACAAAGCACAGTTCTTCCAGTTCCAAATGATGAA 590
Db 481 CCACATGTTCCGGTACATCCGGTGGAACAAAGCACAGTTCTTCCAGTTCCAAATGATGAA 540
QY 591 CATCATCAACGGTGGGGCTCACGCTGACTCCGGTGTGAGTTTCAGGAATTCATGATCGC 650
Db 541 CATCATCAACGGTGGGGCTCACGCTGACTCCGGTGTGAGTTTCAGGAATTCATGATCGC 600
QY 651 TCCAATCGGTGCAGAGACCTTCTCTGAGGCTCTCCGCAACGGCGCGAGGTCTTACCACGC 710
Db 601 TCCAATCGGTGCAGAGACCTTCTCTGAGGCTCTCCGCAACGGCGCGAGGTCTTACCACGC 660
QY 711 ACTGAAGTCCGTTCATCAGAAAGGGCTGTCCACCGGACTTGGCGATGAGGGGGCTT 770
Db 661 ACTGAAGTCCGTTCATCAGAAAGGGCTGTCCACCGGACTTGGCGATGAGGGGGCTT 720
QY 771 CGCTCCCTTCGTCGGCTCCACCGTGAGGCTCTTGACCTTATCGTTAAGGCAATCGAGAA 830
Db 721 CGCTCCCTTCGTCGGCTCCACCGTGAGGCTCTTGACCTTATCGTTAAGGCAATCGAGAA 780
QY 831 GGCTGGCTTCACCCAGGCAAGGACATCGCTCTTGCTCTGGACGTTGCTTCTCTGAGTT 890
Db 781 GGCTGGCTTCACCCAGGCAAGGACATCGCTCTTGCTCTGGACGTTGCTTCTCTGAGTT 840
QY 891 CTTCAAGGACGGACCTTACACCTTCGAAAGTGGCGAGCATCCGCGAGCTGAGATGGCAAA 950
Db 841 CTTCAAGGACGGACCTTACACCTTCGAAAGTGGCGAGCATCCGCGAGCTGAGATGGCAAA 900
QY 951 CGTTTACGCTGAGCTCGTTGACCGCTACCAATCGCTTCCATCGAGGACCCACTGCGAGGA 1010
Db 901 CGTTTACGCTGAGCTCGTTGACCGCTACCAATCGCTTCCATCGAGGACCCACTGCGAGGA 960
QY 1011 AGATGACTGGAGGGTTACACCAACCTCACCGCAACCATCGGCGCAAGGTTCCAGATCGT 1070
Db 961 AGATGACTGGAGGGTTACACCAACCTCACCGCAACCATCGGCGCAAGGTTCCAGATCGT 1020
QY 1071 TGGCGAGCATCTTCTGTCACCAACCTGAGCGCTGAGGAGGGCATCGCTAAGAGGC 1130
Db 1021 TGGCGAGCATCTTCTGTCACCAACCTGAGCGCTGAGGAGGGCATCGCTAAGAGGC 1080
QY 1131 TGCCAACTCCATCTGTTTAAAGTGAACCAAGTACCGTACCGTACCGAGACCTTCCAGCG 1190
Db 1081 TGCCAACTCCATCTGTTTAAAGTGAACCAAGTACCGTACCGTACCGAGACCTTCCAGCG 1140
QY 1191 TGTGCAATGGCTCACCGCGAGGCTACACCTCCATGATGTCCTCCACCGTTCCGGTGAGAC 1250
Db 1141 TGTGCAATGGCTCACCGCGAGGCTACACCTCCATGATGTCCTCCACCGTTCCGGTGAGAC 1200
QY 1251 CGAGGACACACCATTTGCTGACCTCGAGTTGCACTCAACTGFGCCAGATCAAGACTGG 1310
Db 1201 CGAGGACACACCATTTGCTGACCTCGAGTTGCACTCAACTGFGCCAGATCAAGACTGG 1260
QY 1311 TGCTCCAGCAGCTTCCGACCGTGTCCGAAAGTAAACACAGCTTCTCCGATCAGCAGCT 1370
Db 1261 TGCTCCAGCAGCTTCCGACCGTGTCCGAAAGTAAACACAGCTTCTCCGATCAGCAGCT 1320
QY 1371 GCTTGGCGAGCGCGGCTTACCGAGGTGCGACGGATTTCCACGCTTTCCAGGGCTTAAAT 1430
Db 1321 GCTTGGCGAGCGCGGCTTACCGAGGTGCGACGGATTTCCACGCTTTCCAGGGCTTAAAT 1380
QY 1431 AAAAGCGCTTTTCGACGC 1448
Db 1381 AAAAGCGCTTTTCGACGC 1398

```

RESULT 8
 US-09-738-626-1085
 ; Sequence 1085, Application US/09738626
 ; Publication No. US2020197605A1
 ; GENERAL INFORMATION:

```

; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738, 626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 1085
; LENGTH: 1275
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; US-09-738-626-1085

```

```

Query Match      80.7%; Score 1273.4; DB 9; Length 1275;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1274; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 151 GTGGCTGAAATCATGACAGTATTCGCTCCGGAATTTCTCGACTCCCGCGTAACCCAAAC 210
Db 1 GTGGCTGAAATCATGACAGTATTCGCTCCGGAATTTCTCGACTCCCGCGTAACCCAAAC 60
QY 211 GTCGAGCGAGAGGTTTCTCGATGACGGTTCACCGGTGTCGACGGTGTTCATCCGGC 270
Db 61 GTCGAGCGAGAGGTTTCTCGATGACGGTTCACCGGTGTCGACGGTGTTCATCCGGC 120
QY 271 GCATCCACCGGGTCCACGAGGCTCATGAGTGCCTGAGCGGTGCGACGGTGTTCATCCGGC 330
Db 121 GCATCCACCGGGTCCACGAGGCTCATGAGTGCCTGAGCGGTGCGACGGTGTTCATCCGGC 180
QY 331 AAGGGCGTTTGAAGCGAGTTGAACAGTCAACGAGAAATCGGCGACGAGCTCGTGGC 390
Db 181 AAGGGCGTTTGAAGCGAGTTGAACAGTCAACGAGAAATCGGCGACGAGCTCGTGGC 240
QY 391 CTAGAGGCTGACGATCAGCGGCTTCATCGACGAGCAATGATCAAGCTTGTATGGCACCGCC 450
Db 241 CTAGAGGCTGACGATCAGCGGCTTCATCGACGAGCAATGATCAAGCTTGTATGGCACCGCC 300
QY 451 AACAAATCCGCTCGGTGCAACCGCAATCTTGGTGTTCATGCTGCTGTTGCAAAAGGCT 510
Db 301 AACAAATCCGCTCGGTGCAACCGCAATCTTGGTGTTCATGCTGCTGTTGCAAAAGGCT 360
QY 511 GCTGCTGATTCGCGAGGCTCCCACTGTTCCGCTACATCGGTGGACCAACGACAGCTT 570
Db 361 GCTGCTGATTCGCGAGGCTCCCACTGTTCCGCTACATCGGTGGACCAACGACAGCTT 420
QY 571 CTTTCCAGTTCCCAATGATGAACATCATCAACGGTGGCGCTCACGCTGACTCCGGTGTGAC 630
Db 421 CTTTCCAGTTCCCAATGATGAACATCATCAACGGTGGCGCTCACGCTGACTCCGGTGTGAC 480
QY 631 GTTCAGGAATTCATGATCGCTCAATCGGTGAGAGACCTTCTCTGAGGCTCTCCGCAAC 690
Db 481 GTTCAGGAATTCATGATCGCTCAATCGGTGAGAGACCTTCTCTGAGGCTCTCCGCAAC 540
QY 691 GCGCGGAGGCTTACCCACGACCTGAAGTCCGCTCATCAAGGAAAGGGCTGTCCACCGGA 750
Db 541 GCGCGGAGGCTTACCCACGACCTGAAGTCCGCTCATCAAGGAAAGGGCTGTCCACCGGA 600
QY 751 CTTGGCGATGAGGGGGGCTTTCGCTCCTTCGCTCGGCTCCACCGGTGAGGCTCTTGACCTT 810

```

Db	601	CTTGGCGATAGGCGCGCTTCGCTCCTTCGCTCGGCTCCACCCGTGAGGCTTGTGACCTT	660
Qy	811	ATCGTTTAAGGCAATTCGAGAAGCGCTGGCTTCACCCACGACGAAGACATCGCTCTTGCTCTG	870
Db	661	ATCGTTGAGGCAATCGAGNAGCTGGCTTCACCCACGACGAAGACATCGCTCTTGCTCTG	720
Qy	871	GACGTTGCTTCCTCTGAGTCTTCAAGAACGGCACTTACCACTTCGAAGTGGCCAGCAC	930
Db	721	GACGTTGCTTCCTCTGAGTCTTCAAGAACGGCACTTACCACTTCGAAGTGGCCAGCAC	780
Qy	931	TCCGACGCTGAGATGGCAACGTTTACGCTGAGCTCGTTCGACGGTACCCAACTCGTCTCC	990
Db	781	TCCGACGCTGAGATGGCAACGTTTACGCTGAGCTCGTTCGACGGTACCCAACTCGTCTCC	840
Qy	991	ATCGAGGACCCACTGCGAGGAAGATGACTGGGAGGGTTACACCAACCTTCACCGCAACCATC	1050
Db	841	ATCGAGGACCCACTGCGAGGAAGATGACTGGGAGGGTTACACCAACCTTCACCGCAACCATC	900
Qy	1051	GGCGACAAGGTTTTCAGATCGTTGGCGACGACTTCTTCGTTTCGTTCAACACCTGAGCGGCTGAAG	1110
Db	901	GGCGACAAGGTTTTCAGATCGTTGGCGACGACTTCTTCGTTTCGTTCAACACCTGAGCGGCTGAAG	960
Qy	1111	GAGGGCATCGCTAAGNAGGCTGCCAACTCCATCTCTGGTTAAGGTGAACAGATCGGTACC	1170
Db	961	GAGGGCATCGCTAAGNAGGCTGCCAACTCCATCTCTGGTTAAGGTGAACAGATCGGTACC	1020
Qy	1171	CTCACCGAGACCTTTCGACGCTGTGACATGGCTCACCGCGCAGGCTACACCTCCATGATG	1230
Db	1021	CTCACCGAGACCTTTCGACGCTGTGACATGGCTCACCGCGCAGGCTACACCTCCATGATG	1080
Qy	1231	TCCACACGTTCCGGTGTAGACCGAGGACACCAACATTCGCTGACCTTCGAGTTGCACTCAAC	1290
Db	1081	TCCACACGTTTCCGGTGTAGACCGAGGACACCAACATTCGCTGACCTTCGAGTTGCACTCAAC	1140
Qy	1291	TGTGGCCAGATCAAGACTGGTGTCTCCAGCAGTTCCGACCGTGTCCGAAAGTACAAACAG	1350
Db	1141	TGTGGCCAGATCAAGACTGGTGTCTCCAGCAGTTCCGACCGTGTCCGAAAGTACAAACAG	1200
Qy	1351	CTTCTCCGCATTCGAGCAGCTGCTTGGCGACGCGCGGCTCTACGCAGGTTCGACGGCATTC	1410
Db	1201	CTTCTCCGCATTCGAGCAGCTGCTTGGCGACGCGCGGCTCTACGCAGGTTCGACGGCATTC	1260
Qy	1411	CCACGCTTTCAGGGC	1425
Db	1261	CCACGCTTTCAGGGC	1275

RESULT 9

```

US-10-282-122A-1772A
;
; Sequence 1772A, Application US/10282122A
; Publication No. US20040029129A1
;
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR FILING NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23

```



```

QY 811 ATCGTTAAGCAATCGAAGGCTGGCTTCAACCCAGGAAAGACATCGCTCTTGCTCTG 870
Db 661 ATCGTTAAGCAATGAGAAGGCTGGCTTCAACCCAGGCTGCTGACATCGCACTTGCACTC 720
QY 871 GAGGTTGCTTCTCTGAGTTCTTCAAGAGCGGACCTTACCACCTTCAAGAGTGCCAGCAC 930
Db 721 GAGGTTGCTTCTCTGAGTTCTTCAAGAGGATGGAAGTACCACCTTCAAGAGGCGGAGCAC 780
QY 931 TCCGAGCTGAGATGGAACAGTTTACGCTGAGCTGTTGACCGGTACCCAAATCGTCTCC 990
Db 781 ACCGCTGAGGAGATGGCAAGGCTTACGAGCAGCTCATCGCTGAGTACCCAAATGTTTCC 840
QY 991 ATCGAGGACCCATCGCAGGAAGATGACTGGAGGGTTACACCAACCTCACCGCAACATC 1050
Db 841 ATCGAGGACCCATCGCAGGAAGACGACTGGAGGGCTTACACCGCCCTGACCGCGCAATC 900
QY 1051 GCGCAAAAGGTTTCAAGATCGTTGGCGAGCACTTCTTGTCAACCAACCTGAGCGCCTGAAG 1110
Db 901 GGTGACAAGGTTTCAAGATCGTGGGAGCACTTCTTGTCAACCAACCCAGCACGCTCAAG 960
QY 1111 GAGGGATCGCTAAGAGGCTGCCAATCCATCTGTTTAAAGTGAAACGATCGGTACC 1170
Db 961 GAAGGATCGAGAAGAGGCTGCCAATGCTTGTGTAAGGTCAACCAATCGGTACC 1020
QY 1171 CTCACGAGACCTTCGACGCTGTCGACATGGCTCACCGGCGAGCTACACCTCCATGATG 1230
Db 1021 CTGACGAGACCTTCGACGCTGTTGATCTCGCACACCGCAACCGGCTTACCGCACATGATG 1080
QY 1231 TCCACCGGTTCCGGTGAGACCGGAGCACACCAATTTGCTGACCTCGCAGTTGCACTCAAC 1290
Db 1081 TCCACCGGCTCCGGGAGAGCTGAAGACACCAACCAATTTGCTGACCTTGCAGTTCGATGGGC 1140
QY 1291 TGTGGCCAGATCAAGACTGCTGTCGACGAGCTTCCGACGCTGTCGACGAGTTCGCAACCGAG 1350
Db 1141 TGTGGCCAGATCAAGACCGGTCGACGAGCTGTTCCGAGCGCTTTCGCAAGTACAATCAG 1200
QY 1351 CTTCTCGCATCGAGAGCTGCTTGGCGAGCGCGGCTTACGACGAGTTCGACGAGTTC 1410
Db 1201 CTTCTCGCATCGAGAGCTGTTGAGATGACGATGACGAGTTCGACGAGCGCTTACGAGTTC 1260
QY 1411 CCACGCTTTCAGGCG 1425
Db 1261 CCACGCTTTCAGGCG 1275

```

RESULT 10

```

US-10-156-761-1/c
; Sequence 1, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 1
; LENGTH: 9025608
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4187715)

```

```

; OTHER INFORMATION: a, t, c, g, other or unknown
US-10-156-761-1
Query Match 44.2%; Score 696.8; DB 15; Length 9025608;
Best Local Similarity 72.0%; Pred. No. 4.2e-202;
Matches 924; Conservative 0; Mismatches 357; Indels 3; Gaps 1;

QY 140 AGGAGCCACAGTGGCTGAAATCATGACGTATTTCGCTCGGAAATTCGACTCCCGCG 199
Db 4373549 AGGAGATGCTGTGCGCTCCATCGACGTGTCGTAGCCCCGGGAAATTCCTGGACTCCCGAG 4373490
QY 200 GTAACCCAAACCGTCGAGGCGAGAGTTTCTTGATGACCGTTTCCACGCGTGTCCGAGGTG 259
Db 4373489 GCNAACCCACGCTGAGGTGCGGCTCGACGACGCGACGCGGTGTCGCCGCG 4373430
QY 260 TTCCATCCCGGCATCCACCGGCGTCCACGAGCTCATGAGTGTGTCGACGTTG---GCG 316
Db 4373429 TCCGCTCCGGCGCTCCACCGGTGCTTCGAGGCCATCGAGTTCGCGACGCTGACCCCA 4373370
QY 317 ATCGCTTACCTGGGCAAGGCGTTTGAAGGCAAGTTGAAACGTTCAACGAAAGATCGGCG 376
Db 4373369 ACCGTTTACAGGGCAAGGCTGTCGAGAGGCGCTCTCCGCGTTCATCGAGCAGATCGGCC 4373310
QY 377 ACAGCTCCGCTGGCTTAGAGGCTGACGATCAGCGCTCATCGACGAGCAATCATCAAGC 436
Db 4373309 CGGAGCTCGTGGCTACGAGCCACCGAGCAGCGCTGATCGACGAGCGATGTTGACC 4373250
QY 437 TTGATGGCACCGCAACAAAGTCCCGCTGGGTGCAACGCAATCTCTTGGTGTTCATGG 496
Db 4373249 TGGACGCCACCGAACAAAGGCTCGCTCGGGCCCAACGCAATCTCTCGGCTCTCCCTCG 4373190
QY 497 CTGTTGCAAGGCTGCTGCTGATTCCGAGGCTTCCACGTTTCCGCTACATCGGTGGAC 556
Db 4373189 CCGTCGCGACGCGCTCCGAGGCGAGCACTCCCGCTTCTCCCTACCTGGGCGGCC 4373130
QY 557 CAAACGCACACGTTCTTCCAGTTTCCAAATGATGAACATCATCAACGCTGGCGCTCAGCGTG 616
Db 4373129 CGNAACGCGACCTGCTGCCGCTTCCGATGATGNAACATCTGAAACGCGGCTCGCACGCG 4373070
QY 617 ACTCCGGTGTGAGTTTCAAGGAAATTCATGATCGCTCCAAATCGGTGCGAGACCTTCTCTG 676
Db 4373069 ACTCCAAACGTGGACATCCAGGAGTTTCATGATCGCCCCGATCGGCGGAGTCTCTTCCG 4373010
QY 677 AGGCTCTCGCAACGCGGCGGAGGCTTACGACGACTGAAGTCCGTCATCAAGGAAAGG 736
Db 4373009 AGGCCCTGCGCTGGGCGCGGAGGTTTACACACCTTCAAGAAAGGTGCTGAAGACCAAGG 4372950
QY 737 GCCTGTCCACCGACTTGGCGATGAGGCGGCTTTCGCTCTCTTCCGTCGGCTCCACCCGCTG 796
Db 4372949 GCCTGTCCACCGGCTTCCGCGACGAGGCGGCTTCCGCCCGAACCTGGAGTTCGAACCGCG 4372890
QY 797 AGGCTTTGACCTTATCGTTTAAAGGCAATCGAAGGCTGGCTTCAACCCAGGCAAGACA 856
Db 4372889 CCGCGCTCGACCTCATCATCGAGGCCATCAAGCAGGCGGTTTACATCCCGGCGGAGCAGA 4372830
QY 857 TCGCTCTTCTGAGAGTTGCTTCTCTCTGAGTTCTTCAAGGACGCGACTTACACTTCG 916
Db 4372829 TCGCGCTCGCGCTCGACGCTCGCGCTCCGAGTTTCTACAGGACGCGAAGTACGAGTTTCG 4372770
QY 917 AAGGTGGCCAGCACTCCGAGCTGAGATGGCAAAACGTTTACGCTGAGCTGTTGAGCGCT 976
Db 4372769 AGGGCAAGTCCCGCTCGGCCCGCGAGATGACCGAGTACTACGAGGAGCTCGTCTCCGCT 4372710
QY 977 ACCCAATCGTCTCCATCGAGGACCCACTCGAGGAAGATGATCGGAGGTTTACACCAAC 1036
Db 4372709 ACCCGCTCGTCTCCATCGAGGACCCGCTGTACGAGGACGACTTGGGCGCGGTGAAGGTCA 4372650
QY 1037 TCACCGCAACCATCGGCGCAAGGTTTCAAGTTCGTCGCGACGACTTCTTCTGTCACCAAC 1096
Db 4372649 TCACCGCAAGCTGGGCGCAAGGTTCCAGATCGTCCGCGACGACTTCTTCTGTCACCAAC 4372590
QY 1097 CTGAGCGCTTGAAGGAGGCGATCGCTTAAGAGGCTGCCAATCTCCATCTTGTGTTAAGTGA 1156

```


Db 4372589 CGAGGCGCTGGCCCGCGGCATCGAGAGGGCTCCGCCAACGCCCTGCTCGTCAAGGTCA 4372530
QY 1157 ACCAGATCGGTACCTCACCGAGACCTTCAGCGCTGTGCGACATGGCTCACCGCGCAGGCT 1216
Db 4372529 ACCAGATCGGTTCGCTGACCGAGACCTTGAGCGCGCTGAGCTGGCCAGCGCAACGGCT 4372470
QY 1217 ACACCTCCATGATGTCCACCGTTCCGGTGTAGACCGAGGACACCAATTCGTGACTCG 1276
Db 4372469 TCAAGTGCATGATGTCCACCGCTCCGGCGAGACCGAGGACGTCAACATCGCCGACTCG 4372410
QY 1277 CAGTTCACCTCAACTGTGGCCAGATCAAGACTGGTCTCAGCAAGTTCGACCGGTGCG 1336
Db 4372409 CCGTCGCCGTGAATCGCGGTGAGTCAAGACCGCGCGCGCTCGGACCGTGTGCG 4372350
QY 1337 CAAAGTACAACCAAGCTTTCGCCGATCGAGCAGCTGTGTCGCGACGCGCGGTCTACGCGAG 1396
Db 4372349 CCAAGTACAACCAAGCTGCTGCGGATCGAGGAGATCTTCGACGACGCGCGGAGTACGCG 4372290
QY 1397 GTCGACGCGCATTCACCAAGCTTTC 1420
Db 4372289 GCGCTCGCGCTTCCCGCGGTTC 4372266

RESULT 11

US-10-156-761-3520
; Sequence 3520, Application US/10156761
; Publication No. US20030119018A1

GENERAL INFORMATION:

; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 3520
; LENGTH: 1284
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1284)
US-10-156-761-3520

Query Match

Best Local Similarity 44.1%; Score 695.4; DB 15; Length 1284;

Matches 919; Conservative 0; Mismatches 351; Indels 3; Gaps 1;

QY 151 GTGGCTGAATCATGACGATTCCTCGTCGCGAAATTCGACTCCCGCGGTAAACCAACC 210
Db 1 GTGCCGTCCATCATGACGCTGCTAGCGCCGGGAAATCTTGACTCCCGAGGCAACCCACG 60
QY 211 GTGAGGCGAGGTTTCTCGATGACGTTTCCACGCTGTCAGGTGTCCATCCGGC 270
Db 61 GTCAGGTGAGGTGCGGCTCTGACGACGCGAGCAGCGGTGTCGCGCGCTCCGTCGGC 120
QY 271 GCATCCACGGGCTTCCACGAGGCTCATAGCTCGGTGACGGTG---GCGATCGCTACCTG 327
Db 121 GCCTCAACGGTGCCTTCGAGGGCCATCGAGCTCCGCGACGCTGACCCCAACCGTTACCAG 180
QY 328 GGCAGAGGCGTTTGTAGGCGAGTTGAAAGCTCAACGAAGAAATCGGCGACGAGCTCGCT 387
Db 181 GGCAGAGGCTGTCGAGAGGCGCGCTTCGCGCTCATCGAGCAGATCGGCGCGGAGCTCGTC 240

QY 388 GGCCTAGAGGCTGACGATCAGCGCCTCATCGAGCAAGCAATGATCAAGCTTGTATGGCACC 447
Db 241 GGTTCAGACGCGCACCGAGCAGCGCTGATCGACAGCGATGTTGACCTTGGACGCGCAC 300
QY 448 GCCAACAAAGTCCCGCTGGGTGCAAAAGCAATCTTGTGTGTTTCCATGGCTGTGCAAG 507
Db 301 GACAAACAAAGGCTCGCTCGCGCCAAAGGCATCTCTCGGGTCTCTCCCTCGCGCTCGCGCAC 360
QY 508 GGTGCTGCTGATTCGCGACGCGCTCCCACTGTTCCGCTACATCGGTGGACCAAAACGACAC 567
Db 361 GCGCGCTCGAGGCGAGCGACCTCCCGCTCTCCGCTACTCTGGCGGCGCGAAGCGCGCAC 420
QY 568 GTTCTTCCAGTTTCCAAATGATGAACATCATCAACGCTGGCGCTCACGCTGACTCCGCTGT 627
Db 421 CTGCTCCCGCTTCGATGATGAACATCTCAACGCGGCTCGCACGCGGACTCCAAAGTG 480
QY 628 GAGGTTCAAGAAATTCATGATCGCTCCAATCGGTGACAGACCTTCTCTGAGGCTCTCCGC 687
Db 481 GACATCCAGGAGTTTATGATCGCGCCGATCGCGCGGAGTCTTCTCCGAGGCGCTCGCG 540
QY 688 AACGCGCGCGAGGTCTACCAACGACCTGAAGTCCGTCTCAAGGAAAGAGGCTGTCCACC 747
Db 541 TGGGGCGCGAGGTCTACCAACCTCAAGAGGTGCTGAGACCAAGGCGCTGTCCACC 600
QY 748 GGAATTTGGCGATGAGGCGGCTTTCGCTTCCGTCGCTTCCACCGCTGAGGCTCTTTGAC 807
Db 601 GCGCTCGCGCAGAGGCGGCTTTCGCGCCGAACTGGAGTTCGAAACCGCGCGGCTCGAC 660
QY 808 CTTATCGTTAAGCAATCGAAGGCTGGCTTACCCAGGCAAGGCAAGCATCGCTTGTCT 867
Db 661 CTCATCATCGAGGCGCATCAAGCAGGCGGTATACATCCCGGCGAGCAGATTCGCGCTCGCG 720
QY 868 CTGGAGGTTGCTTCTCTGAGTTCTTCAAGGAGCGCACCTTACCCTTTCGAAGGTGGCCAG 927
Db 721 CTCGAGCTCGCGCTCGAGTTCTAAGAGGACGCAAGTACAGTTTCAGAGGCGCAAGTCC 780
QY 928 CACTCCGAGCTGAGATGGCAACGTTTACGCTGAGCTCGTTGACCGGTACCCAAATCGTC 987
Db 781 CGCTCGGCGCGAGATGACCGAGTACTACGAGGAGCTGCTCTCCCGCTACCCGCTCGTC 840
QY 988 TCCATCGAGGACCCACTGCGAGGAGATGATCGGAGGTTTACACCAACCTCACCGCAACC 1047
Db 841 TCCATCGAGGACCCGCTGTACGAGGACGACTGGGCGCGCTGGAGGTCATCACCGCAAG 900
QY 1048 ATCGGCGCAAGGTTCCAGATCGTTGGCGAGCTTCTTGTTCACCAACCTTGAGCGGCTG 1107
Db 901 CTGGGCGCAAGGTCAGATCGTTCGCGGAGACCTCTTGTTCACCAACCGCGAGGCGCTG 960
QY 1108 AAGGAGGCGATCGCTAAGAGGCTGCCAACTCCATCTCGTTAAGGTGAACCGAGATCGGT 1167
Db 961 GCGCGCGCATCGAGGAGGCTCCGCGCAAGCGCTCTGCTCAAGGTCAACAGATCGGT 1020
QY 1168 ACCTCACCAGAGACTTTCAGCGTGTGACATGCTCACCGCGCAGGCTTACCTCATG 1227
Db 1021 TCGCTCACCAGAGACCCCTGGACGCGCTGAGCTGGCGCCAGCGCAACGGCTTCAAGTGCATG 1080
QY 1228 ATGTCCACCGTTTCGCTGAGCCGAGGACACCACTTGTGACCTTCGAGTTGCACTC 1287
Db 1081 ATGTCCACCGCTTCGCGGAGACCGGAGGAGTCACTATCGCGACCTTCGCGCTCGCGGTG 1140
QY 1288 AACTGTGGCGCAGATCAAGACTGCTGCTCCAGCAAGTTCGACCGCTGTGCAAAAGTACAAC 1347
Db 1141 AACTGCGGTGAGATCAAGACCGCGCGCCCGCTCGGACCTGTGCGCAAGTACAAC 1200
QY 1348 CAGCTTCTCCGATTCAGAGCAGTGTGTTGGGAGCGCGGCTGTACGAGGTCGAGCGCA 1407
Db 1201 CAGCTCTCGCATCGAGGAGATCTTCGACGACGCGCGGAGTACGCGCGGCTCGCGG 1260
QY 1408 TTCCCAAGCTTTC 1420
Db 1261 TTCCCCCGGTTC 1273


```

RESULT 14
US-10-080-170-649
; Sequence 649, Application US/10080170
; Publication No. US20030129601A1
; GENERAL INFORMATION:
; APPLICANT: COLE, S.T.
; TITLE OF INVENTION: COMPARATIVE MYCOBACTERIAL GENOMICS AS A TOOL FOR
; TITLE OF INVENTION: IDENTIFYING TARGETS FOR THE DIAGNOSIS, PROPHYLAXIS OR
; TITLE OF INVENTION: TREATMENT OF MYCOBACTERIOSES
; FILE REFERENCE: 03495.0218
; CURRENT APPLICATION NUMBER: US/10/080,170
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: 60/270,123
; PRIOR FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 652
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 649
; LENGTH: 45191
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
US-10-080-170-649

Query Match 40.1%; Score 632.6; DB 15; Length 45191;
Best Local Similarity 68.4%; Pred. No. 3e-183;
Matches 891; Conservative 0; Mismatches 409; Indels 3; Gaps 1;

Qy 140 AGGAGGCCAGGTGGTGAATCATGCACGTATTTCGTCGCGAAATTCCTCGACTCCCGCG 199
Db 1931 AGAGAACCCAGTGGCCGATTATCGAGCAGGTTAGGCGCCGAGAGATCCTCGATTCGCGG 1999
Qy 200 GTAACCCAAACCGTCGAGGCGAGAGTTTCTCTGGATGACGGTTCCCAACGGTGTCCGAGGTG 259

```

19991 GCAACCCGACGGTGGAGTTCGAGTGGCGCTTATCGACGGGACATTCGCCCGGGCGCGG 20050
 260 TTCCATCCGGGCGATCCACCGGCGTCCACGAGGCTCATGAGCTCGGTGACGGTGGCGATC 319
 20051 TGGCGTCCGGGCGCTCCACCGGGGAGCACGAGGCGTTCGAGTTGCGGACGGGCGCGATC 20110
 320 GCTACCTGGCGAAGGGCGTTTGAAGGCGAGTTGAAGAACGTCACGAAGAAATCGGCGACG 379
 20111 GCTACGGCGCAAGGGCTGCAAAAGCGTGCAGGCTGTTCTTGATGAGATCGGCCGG 20170
 380 AGTCTGCTGAGGCTGAGGCTGACGATCAGCGCTTCATCGACGAAGCAATGATCAAGCTTG 439
 20171 CCGTCATCGGACTCAACCGCGACGACGAGGATTTGGTCGACCGAGCGCTGGTGACCTAG 20230
 440 ATGGCAGCGCAACAGTCCGGCTGGGTGGAACGCAATCCTTGGTGTTCATGGCTG 499
 20231 ACGGCACCCCGCAAGTCCCGCTGGGCGGCAACGCGATCTTGGGTGCTCGCTCGCTG 20290
 500 TTGCAAGGCTGTGCTGATTCGCGCAGGCTCCACACTGTTCCGCTACATCGGTGGACCA 559
 20291 TTGCCAAGGCGCGCGGATTCGGCGAGCTGCGGTTGTTCCGTTATGTCGGGGGCGCA 20350
 560 ACGCACAGTCTTCCAGTTCCAAATGATGAACATCAACGCTGGCGCTCACGCTGACT 619
 20351 ACGGCACATCTCGCGGTACCGGTACCGATGATGAACATCTCTCAACGGCGGCACACGCCGATA 20410
 620 CCGGTGTGAGCTTCAGGAATTCATGATCGCTCAATCGGTGAGAGACCTTCTGTAGG 679
 20411 CCGCTGTGACATTTCAAGAGTTTCATGTTGCGGCAATTTGGCGGCGCCAGCTTCGTGAGG 20470
 680 CTCTCCGCAACGGCGCGGAGGCTTACCACGACCTGAAGTCCGTCATCAAGGAAAGGGCC 739
 20471 CGTTGCGCTGGGGCGCTGAGTGTACACGCGCTCAAGTCCGTTCTGTAAGAAAGAGGGC 20530
 740 TGTCCACCGGACTTGGCGATGAGGGCGGCTTCGCTCTCCGTCGGTCCACCGCTGAGG 799
 20531 TGTCCACCGGCTGGGCGACGAAGGGCGCTTCGCGCGGATGTGGCGGCGCACACCGCGG 20590
 800 CTCTTGACCTTATCGTTAAGGCAATCGAAGGCTGGCTTCAACCGAGGCAAGGATCG 859
 20591 CGTTGACCTGTATCAGCGGCGCATCGAGTCGGGCGGCTTGCACCGCGGCGCGAGCTGG 20650
 860 CTCTTGCTTGGAGTGTCTCTGAGTTCTTCAAGGACGGCAC---CTACCACTTCG 916
 20651 CGCTGGCGCTGGAGCGGGCGGCCACCGAGTTCTTACCGAGCGCACGGGTAGCTTCG 20710
 917 AAGTGGCGAGCACTCGCAGCTGAGATGAGCAAGCTTTACGCTGAGCTGTTGACGCGT 976
 20711 AGGGCACCAACCGGTACCGCAGACGAGATGACCGAGTTCTACGGGCGCTGCTCGGCGCT 20770
 977 ACCCAATCGTCTCCATCGAGGACCCACTGCGAGGATGACTGGGAGGTTACACCAAC 1036
 20771 ACCCGTGTGTGATCGATCGAAGACCCACTGTCCGAAGACGATTTGGACGCGTGGGCGCGC 20830
 1037 TCACCGCAACCATCGGCGCAAGAGTTTCAGATCGTTGGCGACGACTTCTTCGTCACCAAC 1096
 20831 TGACGCGCTCGATCGGTGACCGGTTGCAATCGTGGCGACGACATCTTTGTACCAATC 20890
 1097 CTGAGGCGCTGAAGGAGGGGATCGCTAAGAGGCTGCCAATCCATCTCTGTTGAAGTGA 1156
 20891 CCGAGCGGCTCGAGGAGGGGATCGAAACGGGCGCTGGCAATGCTTGTCTGTCGTCGTCG 20950
 1157 ACCAGATCGGTACCTTCACCGAGACCTTCGACGCTGTGACATGGCTCACCGCGAGGCT 1216
 20951 ACCAGATCGGAGCTTGAACGAGACACTCGACGCGGTACGCTGGCTTACCAACGGCGAT 21010
 1217 ACACCTCCATGATGTCCCAACCGTTCCGGTGAGACCGAGGACACCAATTTGCTGACTCG 1276
 21011 ACGCACGATGATCAGTCAACCGGAGTGGGAGACGAGGAGACCAATGATCGCGGACTCG 21070
 1277 CAGTTGCATCACTGTGGCGAGATCAAGACTGGTCTCGAGCACGTTTCGACCGGTGCG 1336
 21071 CCGTGGCCATCGCGACGGGCGAGATCAAGACGGGCGCGCTTGTCTCGAGTGAAGCGGCTCG 21130

1337 CAAAGTACAAACAGGTTCTCCGATCGAGCAGCTGCTTGGCGACGGCGGCTTACGCGAG 1396
 21131 CAAAATACAAACAGGCTGCTCGGATCGAAGAGCGCTTGGCGACGGGCGCGCTACGCGG 21190
 1397 GTCGAGCGCATTTCCACAGCTTTTCAGGGCTTAATAAAGCGCT 1439
 21191 GCGACCTGGCATTTCTCGGTTCCGCTGCGAGACGAATAGGT 21233

RESULT 15
 US-10-080-170-649
 ; Sequence 649, Application US/10080170
 ; Publication No. US20040121322A9
 ; GENERAL INFORMATION:
 ; APPLICANT: COLE, S.T.
 ; TITLE OF INVENTION: COMPARATIVE MYCOBACTERIAL GENOMICS AS A TOOL FOR
 ; IDENTIFYING TARGETS FOR THE DIAGNOSIS, PROPHYLAXIS OR
 ; TREATMENT OF MYCOBACTERIOSES
 ; TITLE OF INVENTION: IDENTIFYING TARGETS FOR THE DIAGNOSIS, PROPHYLAXIS OR
 ; TREATMENT OF MYCOBACTERIOSES
 ; FILE REFERENCE: 03495.0218
 ; CURRENT APPLICATION NUMBER: US/10/080,170
 ; CURRENT FILING DATE: 2002-06-10
 ; PRIOR APPLICATION NUMBER: 60/270,123
 ; PRIOR FILING DATE: 2001-02-22
 ; NUMBER OF SEQ ID NOS: 652
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 649
 ; LENGTH: 45191
 ; TYPE: DNA
 ; ORGANISM: Mycobacterium tuberculosis
 ; US-10-080-170-649

Query Match 40.1%; Score 632.6; DB 19; Length 45191;
 Best Local Similarity 68.4%; Pred. No. 3e-183;
 Matches 891; Conservative 0; Mismatches 409; Indels 3; Gaps 1;

140 AGAGGCCACAGTGGCTGAAATCATGACGATTTCGCTCGGAAATTCGACTCCCGG 199
 19331 AGAGAACCCAGTCCGATTCGACGAGTTAGGCGCCGAGAGATCTCGATTCCCGG 19990
 200 GTAACCCAAACCGTCGAGGCGAGGTTTTCCTGGATGACGCTTCCCAACGGTGTGCGAGGTG 259
 19991 GCAACCCGACGTTGGAGTTCGAGTGGCGCTTATCGACGGGACATTCGCCCGGCGCG 20050
 260 TTCATCCGGCGGATCCACGGGCTCCAGAGGCTCATGAGTGCATGAGTGGCGATC 319
 20051 TGGCGTCGGCGCTCGACCGGGGAGCACGAGGCGTTCGAGTTGCGGACGGCGGATC 20110
 320 GCTACCTGGGCAAGGCGGTTTGAAGGCGAGTTGAAACCTCAACGAAGAAATCGCGGACG 379
 20111 GUTACGGCGCAAGCGGTGCAAAAAGCGCTGCGAGGCTGTTCTTGATGAGATCGGCCCGG 20170
 380 AGCTCGCTGGCTAGAGGCTGACGATCAGCGCTCATCGACGAAGCAATGATCAAGCTTG 439
 20171 CCGTCATCGGACTCAACCGCGACGACGAGTTCGTCGACCGAGCGCTGGTGACCTAG 20230
 440 ATGGCAGCGCAACAGTCCCGCTGGGTGCAAAACGAATCCTTGGTGTTCATGGCTG 499
 20231 ACGGACCCCGGACAGTCCCGCTGGGCGGCAACGCGATCTTGGGTGCTCGCTCGCTG 20290
 500 TTGCAAGGCTGCTGCTGATTCGGAGGCTCCCACTGTTCCGCTACATCGGTGGACCAA 559
 20291 TTGCCAAGGCGGCGGAGTTCGGGAGCTGCGCTGTTGTTCTGTTATGTCGGGGGCGCAA 20350
 560 ACGCACAGCTTCTTCCAGTTCCAATGATGAACATCATCAACGGTGGCGCTCAGCTGACT 619
 20351 ACGCGCACATCTCGCGGTACCGATGATGAACATCTCTCAACGGCGCGCACACGCCGATA 20410
 620 CCGGTGTGACCTTCAGGAATTCATGATCGCTCCAATCGGTGAGAGACCTTCTGTAGG 679
 20411 CCGCTGTGACATTCAGAGATTCATGGTGGCGGCAATTTGGCGGCGCCAGCTTCTGTGAGG 20470
 680 CTCTCCGCAACGGCGCGGAGGTTCTACCAGCACTGAAGTCCGTCATCAAGGAAAGGGCC 739

```
Db      20471  CGTTGGCGCTGGGGCGCTGAGGTGTACCAACGCTCAAGTCGGTCTCAAAAAAGGAGGGC 20530
Qy      740    TGTCCACCGGACTTGGCGATGAGGGGCTTCGCTCCTCCGTCGGCTCCACCCGTGAGG 799
Db      20531  TGTCCACCGGCTTGGGCGAGGAGGGCTTCGCCCGGATGTGGCCGGCACCACCGCGG 20590
Qy      800    CTCTTGACCTTATCGTTAAGCAATCGAAGGCTGGCTTCACCCCAGGCAAGGACATCG 859
Db      20591  CGTTGGACCTGATCAGCCGGGCCATCGAGTCGGGGCTTGCAGCCCGGCCGACGTGG 20650
Qy      860    CTCTTGCTCTGGAGGTTGCTTCTCTGAGTCTTCAAGGACGGAC---CTACCACTTCG 916
Db      20651  CGCTGGCCCTTGGACCGGGCGGCCACCGAGTTCTTACCGACGGCACCGGCTAGCTCTTCG 20710
Qy      917    AAGGTGGCCAGCACTCCGCAGCTGAGATGCAACAGTTTACGCTGAGCTCGTTGACGCGT 976
Db      20711  AGGGACACCCGCTACCGCAGACGAGATGACCGAGTTCTACGGGGGCTCTCGGCGCCT 20770
Qy      977    ACCCAATCGTCTCCATCGAGGACCCACTGCAGGAAGATGACTGGGAGGGTTACACCAACC 1036
Db      20771  ACCCGTGTGTGTGATCGAAGACCCACTGTCCGAAGACGATTGGGACGGCTGGGCCGCGC 20830
Qy      1037  TCACCGCAACCATCGGGCAAGGTTTCAGATCGTTGGCGACGACTTCTTCGTACCAACC 1096
Db      20831  TGACGGCTTCGATCGGTGACCGGGTGCAAAATCGTGGCGACGACATCTTTGTCACCAATC 20890
Qy      1097  CTGAGCGCTGAAGGAGGGCATCGCTAAGAAAGCTGCCAACTCCATCCCTCGTTAAGGTGA 1156
Db      20891  CCGAGCGGCTCGAGGAGGGCATCGAACGGGGCGTGGCAAAATGCGTTGCTCGTCAAGGTGA 20950
Qy      1157  ACCAGATCGGTACCCCTCACCGAGACCTTTCGACGCTGTGACATGGCTCACCGGCGAGGCT 1216
Db      20951  ACCAGATCGGACGTTGACCGAGACACTCGACGGGTTCACGCTGGCTCACCGCGGAT 21010
Qy      1217  ACACCTCCATGATGTCCCAACCGTTCCGGTGAGACCGGAGGACACCAATTTGCTGACCTCG 1276
Db      21011  ACCGCACGATGATCAGTCACCGCAGTGGCGAGACGGAGGACACCATGATCGCCGACCTCG 21070
Qy      1277  CAGTTGCACCTCAACTGTGGCCAGATCAAGACTGGTCTCCAGCACGTTCCGACCGGTGCG 1336
Db      21071  CGGTGGCCATCGGACGGGCGAGATCAAGACGGGCGGCGCTCTCGCAGTGAAGCGGTGCG 21130
Qy      1337  CAAAGTACAAACAGCTTCTCCGATCGAGCAGCTGCTTGGCGACGGCGGCTCTAGCGAG 1396
Db      21131  CAAATACAAACAGCTCTCGGATCGAAGAGGCGCTTGGCGACGGCGCGCTACGCGG 21190
Qy      1397  GTCGACGCGCATTTCCACAGCTTTTCAGGGCTTAATAAAAGCGCT 1439
Db      21191  GCGACCTGGCATTTCTCGGTTCGCGTGGCGTGCGAGACGAAATAGGT 21233
```

Search completed: September 29, 2005, 12:00:10
Job time : 4259.5 secs

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 29, 2005, 02:56:03 ; Search time 28.5 Seconds
(without alignments)
1434.811 Million cell updates/sec

Title: US-10-728-947-4
Perfect score: 2155
Sequence: 1 VAEIMVFAREILDSRGNPT.....QLLGDAGVYAGRSAPFRPQG 425

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1553.5	72.1	429	2 B70623	probable enolase -
2	1502.5	69.7	447	2 G86940	probable enolase [
3	1492.5	66.0	424	2 A97333	enolase (2-phospho
4	1412.5	65.5	426	2 H87462	enolase [imported]
5	1402.5	65.1	429	2 D84094	enolase (2-phospho
6	1391.5	64.5	425	2 A83358	phosphopyruvate hy
7	1377.5	63.9	430	2 B69320	phosphopyruvate hy
8	1359.5	63.1	430	2 AG1381	enolase homolog en
9	1359.5	63.1	430	2 AH1750	enolase homolog en
10	1354.5	62.9	431	2 G96987	enolase [imported]
11	1334.5	61.9	426	2 F70343	enolase - Aquifex
12	1322.5	61.3	422	2 D75251	enolase - Deinococ
13	1320.5	61.3	432	2 S76837	phosphopyruvate hy
14	1318.5	61.2	393	2 AB2752	enolase [imported]
15	1313.5	61.0	430	2 H75022	phosphopyruvate hy
16	1301.5	60.4	428	2 AB1841	phosphopyruvate hy
17	1301.5	60.4	429	2 AC2248	enolase [imported]
18	1301.5	60.4	428	2 A45732	phosphopyruvate hy
19	1298.5	60.3	428	2 D81100	enolase NM81285 [i
20	1297.5	60.2	428	2 F71209	probable phosphogl
21	1283.5	59.6	429	2 G72323	phosphopyruvate hy
22	1280.5	59.4	434	2 T47276	phosphopyruvate hy
23	1279.5	59.4	434	2 A89851	enolase [imported]
24	1278.5	59.3	429	2 H83191	enolase PA3635 [im
25	1272.5	59.0	433	2 D86705	phosphopyruvate hy
26	1258.5	58.4	434	2 D98001	phosphopyruvate hy
27	1256.5	58.3	434	2 E95130	enolase [imported]
28	1249.5	58.0	433	2 H70141	enolase (eno) homo
29	1241.5	57.6	427	2 A64329	phosphopyruvate hy

30	1232.5	57.2	432	2 F71278	probable enolase (
31	1229.5	57.0	432	1 NOEC	phosphopyruvate hy
32	1229.5	57.0	432	2 G91083	enolase [imported]
33	1226.5	56.9	432	2 H85928	enolase [imported]
34	1224.5	56.8	432	2 AC0859	enolase [imported]
35	1222.5	56.7	431	2 AB0410	phosphopyruvate hy
36	1218.5	56.5	436	2 E64103	phosphopyruvate hy
37	1211.5	56.2	461	2 F82700	enolase XF1291 [im
38	1204.5	55.9	422	2 D86659	phosphopyruvate hy
39	1178.5	54.7	414	2 E81264	phosphopyruvate hy
40	1174.5	54.5	433	2 E82076	enolase VC2447 [im
41	1168.5	54.2	428	2 A72034	enolase - Chlamydo
42	1165.5	54.1	428	2 G81504	enolase CP1071 [im
43	1165.5	54.1	428	2 F86590	enolase [imported]
44	1155.5	53.6	424	2 E81654	enolase TC0876 [im
45	1155.5	53.6	444	2 T12341	phosphopyruvate hy

ALIGNMENTS

RESULT 1

B70623

Probable enolase - Mycobacterium tuberculosis (strain H37RV)

C:Species: Mycobacterium tuberculosis

C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 16-Aug-2004

C:Accession: B70623

R:Coile, S.T.; Brosch, R.; Garnier, T.; Parkhill, J.; Garnier, C.; Harris, D.; Gordon, S.

R:Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.

Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

Nature 393, 537-544, 1998

A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome

A:Reference number: A70500; MUID:98295987; PMID:9634230

A:Accession: B70623

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-429 <COL>

A:Cross-references: UNIPROT:P96377; GB:292539; GB:AL123456; NID:g3261714; PIDN:CAB06856.

A:Experimental source: strain H37Rv

C:Genetics:

A:Gene: eno

C:Superfamily: Enolase

Query Match	72.1%	Score	1553.5;	DB 2;	Length	429;
Best Local Similarity	73.2%	Pred.	No. 1.3e-93;			
Matches	308;	Conservative	42;	Mismatches	70;	Indels 1; Gaps 1;
QY	4	IMHVFAREILDSRGNPTVEAEVFLDDGSHGVAGVPSGASTGVHAEHLRDRGDRYLKGV	63			
Db	4	IEQVRAEILDSRGNPTVEAEVFLDDGSHGVAGVPSGASTGVHAEHLRDRGDRYLKGV	63			
QY	64	LKAVENNEIEGDELGLADRDORLIDEAMIKLDGTANKSRIGANAILGVSMVAKAAAD	123			
Db	64	QKAVQVLDEIGFPAVIGLNADDQRLVDQALVDLDGTPDKSRLLGNAILGVSLAVAKAAAD	123			
QY	124	SAGLPLFRYVGGPNAHLVPMNIIINGGAHDSGVDOEFMIAPICAEFTSEALRNGAE	183			
Db	124	SAELPLFRYVGGPNAHLVPMNIIINGGAHDSGVDOEFMIAPICAEFTSEALRNGAE	183			
QY	184	VYHALKSVIKERKGLSTGLDEGGFAPSVGSTRALDLIVKAIKAGFTPGKDIALLADVA	243			
Db	184	VYHALKSVLKKEGLSTGLDEGGFAPDPVAGTTAAALDLISRAIESAGLRPGADVALALDAA	243			
QY	244	SSEFFKDGKT-YHPEGGQHSAAEWANVYABLVDAVPIVSIEDPLQEDDWEQYTWLTATIGD	302			
Db	244	ATEFFTDGTGVYFEGTTRTAQWTEFYAGLLGAYPLVSIEDPLSDDWDGMAALTASIGD	303			
QY	303	KVOIVGDDDFVTNPERLKEGIKAKKANSILVKVNOIGTLTETFDVDMAHACGYTSMWSH	362			
Db	304	RVOIVGDDDFVTNPERLKEGIKAKKANSILVKVNOIGTLTETFDVDMAHACGYTSMWSH	363			
QY	363	RSGETEDTTIADLAVALNCCQIKTGAPARSDRVAKYNQLLRIBQLLGDAGVYAGRSAPFR	422			

Db 364 RSGTETMTADLAVALGSGQIKTGAPARSERVAKYNQLRIEBALGDAAARYAGDLAFPR 423
Qy 423 F 423
Db 424 F 424

RESULT 2
G86940
probable enolase [imported] - Mycobacterium leprae
C:Species: Mycobacterium leprae
C>Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 10-May-2001
C:Accession: G86940
R.; Cole, S.T.; Eiglmeier, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; Ho
eum, M.A.; Rutherford, K.M.
Nature 409, 1007-1011, 2001
A:Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.; S
A:Title: Massive gene decay in the leprosy bacillus.
A:Reference number: A86909; MUID:21128732; PMID:11234002
A:Accession: G86940
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-447 <STO>
A:Cross-references: GB:AL450380; NID:gl3092594; PIDN:CAC29763.1; GSPDB:GN00147
C:Genetics:
A:Gene: eno
C:Superfamily: enolase

Query Match 69.7%; Score 1502.5; DB 2; Length 447;
Best Local Similarity 70.8%; Pred. No. 2.8e-90;
Matches 300; Conservative 47; Mismatches 76; Indels 1; Gaps 1;

Qy 1 VAEMHVFAREILDSRGNTPTVEAEVFLDDGSHGVAGVPSGASTGVHAEHLRDGDRYL 60
Db 19 VPVIEQVGAREILDSRGNTPTVEAEVFLDDGSHGVAGVPSGASTGVHAEHLRDGDRY 78
Qy 61 KGVKAVENVNBEIGDELAGLEADDORLIDEAMIKLDGTANKSRILGNAILGVSMVAKA 120
Db 79 KGVRAVDVLDIEGPVVIIGNANDORLIDQELLDLDTGTPDKSRILGNAILGVSLAVAKA 138
Qy 121 AADSAGLPLFRYIGGNNAHLVPVPMNIIINGGAHDSGVVDVQEFMIAPITGAETFSBALRN 180
Db 139 AADSAGLPLFRYIGGNNAHLVPVPMNIIINGGAHDSGVVDVQEFMIAPITGAETFSBALRN 198
Qy 181 GARVYHALKSVIIEKGLSTGLDEGGFAPSGVSTREALDLIVKAIEKAGTTPGKDIALAL 240
Db 199 GARVYHALKSVIIEKGLSTGLDEGGFAPSGVSTREALDLIVKAIEKAGTTPGKDIALAL 258
Qy 241 DVASSEFFKDDGT-YHPEGGGSHAAEMANVYAEVLVDAPVIVSIEDPLQEDDMEGYTNLTAT 299
Db 259 DAAATFYTDCIGYHPEGTHTDQWTEFFADLGLSPLVSIEDPLQEDDMEGYTNLTAT 318
Qy 300 IGDKVQIVGDDFFVTNPERLKEGIAKKAANSILVKVNIQIGTLTETFDVADMAHAGYTS 359
Db 319 IGDKVQIVGDDFFVTNPERLKEGIAKKAANSILVKVNIQIGTLTETFDVADMAHAGYTS 378
Qy 360 MHSRGTEDTTADLAVALNCGQIKTGAPARSDRVAKYNQLRIEQLLDAGVYAGRSA 419
Db 379 ISHRSGETEDTMTADLAVALNCGQIKTGAPARSDRVAKYNQLRIEQLLDAGVYAGRSA 438
Qy 420 FPRF 423
Db 439 FLRY 442

RESULT 3
A97533
enolase (2-phosphoglycerate dehydratase) (2-phospho-D-glycerate hydro-lyase) [imported]
C:Species: Aerobacterium tumefaciens
C>Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 16-Aug-2004
C:Accession: A97533

R.; Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorollo, B.; Goldman,
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A:Reference number: A97359; MUID:21608551; PMID:11743194
A:Accession: A97533
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-424 <KUR>
A:Cross-references: UNIPROT:Q8UPH1; GB:AE007869; PIDN:AAK87218.1; PID:gl5156500; GSPDB:G
C:Genetics:
A:Gene: AGR_C_2631
A:Map position: circular chromosome
C:Superfamily: enolase

Query Match 66.0%; Score 1422; DB 2; Length 424;
Best Local Similarity 66.2%; Pred. No. 4.5e-85;
Matches 276; Conservative 58; Mismatches 81; Indels 2; Gaps 1;

Qy 4 IMHVFAREILDSRGNTPTVEAEVFLDDGSHGVAGVPSGASTGVHAEHLRDGDRYLKGV 63
Db 4 ITDIIAREILDSRGNTPTVEAEVFLDDGSHGVAGVPSGASTGVHAEHLRDGDRYLKGV 63
Qy 64 LKAVENVNBEIGDELAGLEADDORLIDEAMIKLDGTANKSRILGNAILGVSMVAKAAAD 123
Db 64 EKAVEAVNTEIFDAIGCFDRENIQIDQMIALDGTNPKSRILGNAILGVSLAIAKAAAE 123
Qy 124 SAGLPLFRYIGGNNAHLVPVPMNIIINGGAHDSGVVDVQEFMIAPITGAETFSBALNGAE 183
Db 124 ASGLPLRYVGGPNNAHLVPVPMNIIINGGAHDSGVVDVQEFMIAPITGAETFSBALNGAE 183
Qy 184 VYHALKSVIIEKGLSTGLDEGGFAPSGVSTREALDLIVKAIEKAGTTPGKDIALDVA 243
Db 184 VFHTLKKELSAQCHNTNVGDEGGFAPSGVSTREALDLIVKAIEKAGTTPGKDIALDVA 243
Qy 244 SSEFFKDDGT-YHPEGGGSHAAEMANVYAEVLVDAPVIVSIEDPLQEDDMEGYTNLTAT 301
Db 244 STEFFKDDGT-YHPEGGGSHAAEMANVYAEVLVDAPVIVSIEDPLQEDDMEGYTNLTAT 303
Qy 302 DKVQIVGDDFFVTNPERLKEGIAKKAANSILVKVNIQIGTLTETFDVADMAHAGYTSMS 361
Db 304 NKQQLVGGDLFTVNSARLDGIMKGVANSILVKVNIQIGTLTETFDVADMAHAGYTSMS 363
Qy 362 HRSGETEDTMTADLAVALNCGQIKTGAPARSDRVAKYNQLRIEQLLDAGVYAGRS 418
Db 364 HRSGETEDTMTADLAVALNCGQIKTGAPARSDRVAKYNQLRIEQLLDAGVYAGRS 420

RESULT 4
H87462
enolase [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C>Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C:Accession: H87462
R.; Nierman, W.C.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: A87249; MUID:21173698; PMID:11259647
A:Accession: H87462
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-426 <STO>
A:Cross-references: UNIPROT:Q9A7J9; GB:AE005673; NID:gl3423140; PIDN:AAK23700.1; GSPDB:G
C:Genetics:
A:Gene: CC1724
C:Superfamily: enolase

Query Match 65.5%; Score 1412.5; DB 2; Length 426;
Best Local Similarity 66.4%; Pred. No. 1.9e-84;
Matches 280; Conservative 59; Mismatches 80; Indels 3; Gaps 2;

QY 1 VAEIMHVFAREILDSRGNTPTVEAEVFLDDGSHGVAGVPSGASTGVHEAHELDRDGD-RYL 59
Db 1 MTEIVDIIAREILDSRGNTPTVEVDVILEDCGAFGRAAVPSGASTGAHEANEKRDGDKARYL 60
QY 60 GKGVLAKEVNVNEEIGDELAGLEADDOORLIDEAMIKLDGTANKSRILGANAILGVSMVAK 119
Db 61 GKGVQAVADVNGEIEFDALSGVDAEQRRVDNLMIELDGTNPKARILGANAILGVSLATAK 120
QY 120 AAADSAGLPLFRYIGGPNNAHVLVPMWNIINGGAHADSGVDVQEFMIAPIGASTFSEALR 179
Db 121 AAESAGLPLKYVGVGNARVLVPTWNIINGGAHADNPIDIOEFMILPTGAKDFREGLR 180
QY 180 NGAEVHALKSVIKEXGLSTGLDGEFGFAPSVGSTREALDLIVKAIEKAGFTPGKDIALA 239
Db 181 MGAEIFHALKALKADAGHNTNVGDEGFGFAPNLASAEALDFIVKAGEKAGYKAGDFFVLG 240
QY 240 LDVASSEFFKDGTYHFEQGHHS--AAEMANVYAEILDVDAPIVSIEDPLQEDDWEQYTNLT 297
Db 241 LDVASTEFFFFKNGKYELEGESLDPAAMVDYLAGLVAKFFPILTIEDGMAEDDDFDGKKLLT 300
QY 298 ATIGDKVQIVGDDFFVTNPERLKEGIAKKAANSILVKVNOIGTLTETFDVDMHARAGYT 357
Db 301 DTGLGKQVGVDDLLFTNPKRQWGLDKGLANSILVKVNOIGTLSETIDAVELAHRHGYT 360
QY 358 SMHSRSGEETDTIADLAVALNCGQIKTGAPARSDRVAKYNQQLRIEQLLGDAGVYAGR 417
Db 361 SVMHSRSGEETDTIADLAVALNCGQIKTGLARSRTAKYNQQLRIEQLLGDAGVYAGR 420
QY 418 SA 419
Db 421 AA 422
RESULT 5
D84094
enolase (2-phosphoglycerate dehydratase) eno [imported] - Bacillus halodurans (strain C-
C:Species: Bacillus halodurans
C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 16-Aug-2004
C:Accession: D84094
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, F.; Hirai
Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A;Reference number: A83650; MUID:20512582; PMID:11058132
A;Accession: D84094
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-429 <STO>
A;Cross-references: UNIPROT:Q9K717; GB:AP001519; GB:BA000004; NID:g10176109; PIDN:BA072
A;Experimental source: strain C-125
C;Genetics:
A;Gene: eno
C;Superfamily: Enolase
Query Match 65.1%; Score 1402; DB 2; Length 429;
Best Local Similarity 65.8%; Pred. No. 9.1e-84;
Matches 277; Conservative 60; Mismatches 80; Indels 4; Gaps 2;
QY 4 IMHVFAREILDSRGNTPTVEAEVFLDDGSHGVAGVPSGASTGVHEAHELDRDGGDRYLGKGV 63
Db 4 ITDVAREVLDSRGNTPTVEEVVLESYGAMGRALVPSGASTGTEAEVELRDGGERFLGKGV 63
QY 64 LKAVERNNEEIGDELAGLEADDOORLIDEAMIKLDGTANKSRILGANAILGVSMVAKAAAD 123
Db 64 LKAVERNNEEIVAPLIGFDALDQIDQHMIELDGTENKGLGANAILGVSMVAVARAAN 123
QY 124 SAGLPLFRYIGGPNNAHVLVPMWNIINGGAHADSGVDVQEFMIAPICAEFTFSEALRNGAE 183
Db 124 ALDPLVYVLLGGFNAKQLVPMWNIINGGEHADNNVDIQEFMIMPVGAESFKEALRTGTE 183
QY 184 VYHALKSVIKEGLSTGLDGEFGFAPSVGSTREALDLIVKAIEKAGFTPGKDTALALDVA 243
Db 184 IFHSLKVLKSGKYNATVGDGEGFAPNLSSNEEALQTIIEBAEQAGYTPGEQVKLAMDA 243

QY 244 SSEFF--KDGTYHF--EGGQHSAAEMANVYAEILDVDAPIVSIEDPLQEDDWEQYTNLTAT 299
Db 244 SSELNKGDKYHLSGEGVLSSEEMVAFYEELVAKYPIISIEDGLDENDWECHKMLTDR 303
QY 300 IGDKVQIVGDDFFVTNPERLKEGIAKKAANSILVKVNOIGTLTETFDVDMHARAGYTSM 359
Db 304 LGDKVQLVGVDDLLFTNKKLAQIEQGVNSILIKVNOIGTLTETFDATBEMAKRAGYTAV 363
QY 360 MHSRSGEETDTIADLAVALNCGQIKTGAPARSDRVAKYNQQLRIEQLLGDAGVYAGRSA 419
Db 364 IHSRSGEETDTIADLAVALNCGQIKTGAPARSDRVAKYNQQLRIEQLLGDAGVYAGRS 423
QY 420 F 420
Db 424 F 424
RESULT 6
AE3358
phosphopyruvate hydratase (EC 4.2.1.11) [imported] - Brucella melitensis (strain 16M)
C:Species: Brucella melitensis
C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 16-Aug-2004
C:Accession: AE3358
R;DelVecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova, I.
.; Mazur, M.; Goltzman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letessier
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A;Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis;
A;Reference number: AD3252; PMID:11756688
A;Accession: AE3358
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-425 <KUR>
A;Cross-references: UNIPROT:Q8YHF0; GB:AE008917; PIDN:AAL52032.1; PID:g17982798; GSPDB:GSPDB:
A;Experimental source: strain 16M
C;Genetics:
A;Gene: BMEI0851
A;Map position: I
C;Superfamily: Enolase
C;Keywords: Carbon-oxygen lyase; hydro-lyase
Query Match 64.5%; Score 1391; DB 2; Length 425;
Best Local Similarity 65.6%; Pred. No. 4.6e-83;
Matches 274; Conservative 58; Mismatches 84; Indels 2; Gaps 1;
QY 4 IMHVFAREILDSRGNTPTVEAEVFLDDGSHGVAGVPSGASTGVHEAHELDRDGGDRYLGKGV 63
Db 4 IIDIVGREILDSRGNTPTVEVDVLEDSFGRAAVPSGASTGAHEVELRDGGSRYLGKGV 63
QY 64 LKAVERNNEEIGDELAGLEADDOORLIDEAMIKLDGTANKSRILGANAILGVSMVAKAAAD 123
Db 64 EKAVEVNGKIFDAIAGMDAESQLLIDQLDLDGSANKGNLGNAILGVSLAVAKAAQ 123
QY 124 SAGLPLFRYIGGPNNAHVLVPMWNIINGGAHADSGVDVQEFMIAPICAEFTFSEALRNGAE 183
Db 124 ASGLPLRYVVGNTNAHVLVPMWNIINGGAHADNPIDIOEFMILPVGATSIKREAVRYGSE 183
QY 184 VYHALKSVIKEGLSTGLDGEFGFAPSVGSTREALDLIVKAIEKAGFTPGKDTALALDVA 243
Db 184 VFHTLKKRLKLDAGHNTNVGDEGFGFAPNLKNAQAALDFIMESIEKAGFKPGEDIALGLDCA 243
QY 244 SSEFFKDGTYHFEQGHQ--HSAEMANVYAEILDVDAPIVSIEDPLQEDDWEQYTNLTATIG 301
Db 244 ATEFFKDGNYVEGERKTRDPAQAKYLAKLASDYPIVITIEDGMAEDDWEKWLTDLIG 303
QY 302 DKVQIVGDDFFVTNPERLKEGIAKKAANSILVKVNOIGTLTETFDVDMHARAGYTSMMS 361
Db 304 NKQQLVGVDDLLFTNNSARLDGIRLVANSILVKVNOIGLSLSETLDVETAHAKAGYTAAMS 363
QY 362 HRSGETEDTTIADLAVALNCGQIKTGAPARSDRVAKYNQQLRIEQLLGDAGVYAGRSA 419
Db 364 HRSGETEDTTIADLAVALNCGQIKTGLARSRTAKYNQQLRIEQLLGDAGVYAGRSA 421

RESULT 7
B69620
phosphorylase hydratase (EC 4.2.1.11) [validated] - Bacillus subtilis
N:Alternate names: enolase
C:Species: Bacillus subtilis
C:Date: 05-Dec-1997 #sequence revision 05-Dec-1997 #text_change 16-Aug-2004
C:Accession: B69620; I40025; A41835
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chd
A.; Ehrlich, S.D.; Emmeron, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler
lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maue
Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon
A:Authors: Schleich, S.; Schroter, R.; Scoffone, P.; Sekiguchi, J.; Sekowska, A.; Seron
skeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpetra, P.; Tognoni, A.; Tosato, V.; Uchiyama
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A:Reference number: A69580; MUID:98044033; PMID:9384377
A:Accession: B69620
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-430 <KUN>
A:Cross-references: UNIPROT:P37869; GB:Z99121; GB:AL009126; NID:G2635827; PIDN:CAB15395.
A:Experimental source: strain 168
R:Leyva-Vazquez, M.A.; Setlow, P.
J. Bacteriol. 176, 3903-3910, 1994
A:Title: Cloning and nucleotide sequences of the genes encoding triose phosphate isomera
A:Reference number: I40022; MUID:94292408; PMID:8021172
A:Accession: I40025
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-44, 'Q', 46-430 <LEY>
A:Cross-references: GB:L29475; NID:G460256; PIDN:AAA21681.1; PID:G460259
R:Mitchell, C.; Morris, P.W.; Vary, J.C.
J. Bacteriol. 174, 2474-2477, 1992
A:Title: Identification of proteins phosphorylated by ATP during sporulation of Bacillus
A:Reference number: A41835; MUID:92210489; PMID:1556067
A:Accession: A41835
A:Molecule type: protein
A:Residues: 2-10 <MIT>
A:Note: this protein is phosphorylated during exponential growth and stationary phase
C:Genetics:
A:Gene: eno
C:Function:
A:Description: EC 4.2.1.11 [validated, MUID:94292408]
C:Superfamily: Enolase
C:Keywords: carbon-oxygen lyase; hydro-lyase; phosphoprotein

Query Match 63.9%; Score 1377.5; DB 2; Length 430;
Best Local Similarity 64.2%; Pred. No. 3.6e-82;
Matches 271; Conservative 62; Mismatches 84; Indels 5; Gaps 3;
QY 4 IMHVFAREILDSRGNTPTVEAEVFLDDGSHGVAGVPSGAGTVGHEAHELRDGG-DRVLGKG 62
DB 4 IVDVYAREVLDSRGNTPTVEVEVETGAFGRALVPSGASTGGEYAEVLELDGDKDRYLKG 63
QY 63 VLKAVENNEEIGDELGLAGLADDDQRLIDEAMIKLDGTANKSRGLGANAILGVSMVAKAAA 122
DB 64 VLTAVNNNEIIPPELLGPDVTQNAIDQLLIEDGTENKGLGANAILGVSMACARAA 123
QY 123 DSAGLPLFRYIGGPNHVLVPMNMINGGAHDSGVQVDFEPMIPIGAETSEALRNCA 182
DB 124 DFLQIFLYQLGGFNSKTLFVPMNMINGGEHADNNVDIQEEMIMVPGAPNFREALRMCA 183
QY 183 EVYHALKSVIKKGLSTGLDGGEGFAPSVGSTREALDLIVKAEKAGFTPGKDIALALDV 242
DB 184 QIFHLSKSVLSAKGLTAVGDEGGFAPNLGSNEEALQTVAELEKAGFPGEVKLMDA 243

QY 243 ASSEFF--KDGTYHF--EGQHSAAEMANVYAEVLVDAYPIVSIEDPQEDDWEGYTNLTA 298
DB 244 ASSEFYNKEDGKYLSEGEGVVKTSABWVDWYELVSKYPIISIEDGLDENDWEGHKLLE 303
QY 299 TIGDKVQIVGDDDFVTNPNRLKEGIAKKAANSILVKVNQIGTLTETTFDAVDMHARAGYTS 358
DB 304 RLQKQVQLVGDDDFVTNPKLSEGIKNGVNSILIKVNQIGTLTETTFDAIEMAKRAGYTA 363
QY 359 MMSHRSGETEDTTIADLAVALNCQIKTGAPARSDRVAKYNQLLRLEQLLGDAGVYAGRS 418
DB 364 VISHRSGETEDSTIADIATNAGQIKTGAPSRDTRVAKYNQLLRLEDQLAETAQYHGIN 423
QY 419 AF 420
DB 424 SF 425
RESULT 8
AG1381
enolase homolog eno [imported] - Listeria monocytogenes (strain EGD-e)
C:Species: Listeria monocytogenes
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 16-Aug-2004
C:Accession: AG1381
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker,
D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.,
Science 294, 849-852, 2001
A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Maqueno, E.; Maitournam, A.; Mat
ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Weiland,
A:Title: Comparative genomics of Listeria species
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AG1381
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-430 <GLA>
A:Cross-references: UNIPROT:O92813; GB:NC_003210; PIDN:CAD00533.1; PID:G16411943; GSPDB:
A:Experimental source: strain EGD-e
C:Genetics:
A:Gene: eno
C:Superfamily: Enolase
Query Match 63.1%; Score 1359.5; DB 2; Length 430;
Best Local Similarity 63.7%; Pred. No. 5.3e-81;
Matches 269; Conservative 60; Mismatches 88; Indels 5; Gaps 3;
QY 4 IMHVFAREILDSRGNTPTVEAEVFLDDGSHGVAGVPSGASTGVHEAHELRDGG-RYLGKG 62
DB 4 ITEVYAREVLDSRGNTPTVEVEVTEAGAFGRALVPSGASTGGEYAEVLELDGDKARYLGKG 63
QY 63 VLKAVENNEEIGDELGLAGLADDDQRLIDEAMIKLDGTANKSRGLGANAILGVSMVAKAAA 122
DB 64 VLKAVENNDIADKIIGFDVTDQIGDKAMTELDGTPNKGKLGANAILGVSLAARAA 123
QY 123 DSAGLPLFRYIGGPNHVLVPMNMINGGAHDSGVQVDFEPMIPIGAETSEALRNCA 182
DB 124 DELGVHLYEYLGVGNGKVLVPMNMINGGEHADNNVDVQEFMVPVGPAPNKEALRMCA 183
QY 183 EVYHALKSVIKKGLSTGLDGGEGFAPSVGSTREALDLIVKAEKAGFTPGKDIALALDV 242
DB 184 ETLHALKAVLKGKGLNTGVDEGGFAPNLKSNEEALETIMQAIKDAKYGKPGGEVKLMDA 243
QY 243 ASSEFF--KDGTYHF--EGQHSAAEMANVYAEVLVDAYPIVSIEDPQEDDWEGYTNLTA 298
DB 244 ASSEFYNETGKYLKGEVTRTSEMTWTVEEMTKYPIISIEDGLDENDWDGFKLLTE 303
QY 299 TIGDKVQIVGDDDFVTNPNRLKEGIAKKAANSILVKVNQIGTLTETTFDAVDMHARAGYTS 358
DB 304 RIGDRVQLVGDDLFVTNPKLKEGIEKGIANSILIKVNQIGTLTETTLDAISMAKAGYTA 363
QY 359 MMSHRSGETEDTTIADLAVALNCQIKTGAPARSDRVAKYNQLLRLEQLLGDAGVYAGRS 418
DB 364 VISHRSGETEDSTIADIATNAGQIKTGAPTRDTRVAKYNQLLRLEDLADLAETAEYHGD 423

Qy 419 AF 420

Db 424 TF 425

RESULT 9

AH1750

enolase homolog eno [imported] - Listeria innocua (strain Clp11262)

C;Species: Listeria innocua

C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 16-Aug-2004

C;Accession: AH1750

R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker

D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Duseurget, O.; Entian, K.D.; Fsihi, H.

S.; Jones, L.M.; Karst, U.

Science 294, 849-852, 2001

A;Authors: Krest, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma

ok, C.; Schlueter, T.; Simeos, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend,

A;Title: Comparative genomics of Listeria species.

A;Reference number: AB1077; MUID:21537279; PMID:11679669

A;Accession: AH1750

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-430 <GLA>

A;Cross-references: UNIPROT:Q92813; GB:AL592022; PIDN:CAC97776.1; PID:g16415071; GSPDB:G

A;Experimental source: strain Clp11262

C;Genetics:

A;Gene: eno

C;Superfamily: Enolase

Query Match

Best Local Similarity 63.1%; Score 1359.5; DB 2; Length 430;

Matches 269; Conservative 60; Mismatches 88; Indels 5; Gaps 3;

Qy 4 IMHVFAEILDSRGNPTVEAEVFLDDGSHGVAGVPFGASTGVVHEAHELRDGGD-RYLKKG 62

Db 4 ITEVAREVLDSRGNPTVEAEVFLDDGSHGVAGVPFGASTGVVHEAHELRDGGD-RYLKKG 63

Qy 63 VLKAVENNEEIGDELADDDORLIDEAMIKLDTANKSRGLGANAILGVSMVAKAAA 122

Db 64 VLKAVENNDIADKIIGFDVTQIGIDKAMIELDTGPNKGKLGANAILGVSLAARAAA 123

Qy 123 DSAGLPLFRYIGGNPAHVLVPMNMNIINGGAHDSGVDPVQEFMIAPIGATTFSEALRNGA 182

Db 124 DELGVHLYELGVNGKVLVPMNMNIINGGAHDSGVDPVQEFMIAPIGATTFSEALRNGA 183

Qy 183 EVYHALKSVLKEGLSTGLDGGFAPSGVSTREALDLIVKAEKAGFTPGKDIALDV 242

Db 184 EILHALKAVLKGKLGNTGVGDEGGFAPNLKSNSEALETIMQAIKDAKYGKPEEVKLAMD 243

Qy 243 ASSEFF--KDGTYHF--EGGQHSAAEMANVYAEVLVDAYPIVSIEDDPLQEDDWEYTNLTA 298

Db 244 ASSEFFNRETGKVELKGEVTRISEEMVTVYEEMITKYPISIEDGLDENDWGFLLTE 303

Qy 299 TIGDKVQIVGDDPFTVTPERLKEGIAKANSILVKVNOIGTTLTETFPDVMMAHRAGYTS 358

Db 304 RIGDRVQLVGDLLFVNTTKLKEGIEKGIANSILVKVNOIGTTLTETFPDVMMAHRAGYTA 363

Qy 359 MSHRSGETEDTTIADLAVALNCGQIKTGAPARSDRVAKYNQLLRIBOLLGDAGVYAGRS 418

Db 364 VISHRSGETEDTTIADLAVATNAAGQIKTGAPTRDRVAKYNQLLRIBOLLGDAGVYAGRS 423

Qy 419 AF 420

Db 424 TF 425

RESULT 10

G96987

enolase [imported] - Clostridium acetobutylicum

C;Species: Clostridium acetobutylicum

C;Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 09-Jul-2004

C;Accession: G96987

R;Nolling, J.; Breton, G.; Omeichenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,

.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.

J. Bacteriol. 183, 4823-4838, 2001

A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clo

A;Reference number: A96900; MUID:21359325; PMID:21359325

A;Accession: G96987

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-431 <KUR>

A;Cross-references: UNIPROT:Q97L52; GB:AE001437; PIDN:AAK78690.1; PID:g15023593; GSPDB:G

A;Experimental source: Clostridium acetobutylicum ATCC824

C;Genetics:

A;Gene: CAC0713

C;Superfamily: enolase

Query Match 62.9%; Score 1354.5; DB 2; Length 431;

Best Local Similarity 62.5%; Pred. No. 1.1e-80;

Matches 263; Conservative 73; Mismatches 82; Indels 3; Gaps 2;

Qy 3 EIMHVFAEILDSRGNPTVEAEVFLDDGSHGVAGVPFGASTGVVHEAHELRDGGD-RYLKKG 61

Db 6 EIVDMARQILDSRANPTVEAEVFLDDGSHGVAGVPFGASTGVVHEAHELRDNDKQYLGK 65

Qy 62 GVLKAVENNEEIGDELADDDORLIDEAMIKLDTANKSRGLGANAILGVSMVAKAAA 121

Db 66 SVLNADVNVNTEITATLIGNVDFDTLIDQTMLEIDGTENKSKLGANAILGVSLAVAAA 125

Qy 122 ADSAGLPLFRYIGGNPAHVLVPMNMNIINGGAHDSGVDPVQEFMIAPIGATTFSEALRNG 181

Db 126 AEVLGSLQYLGGVNAKVLVPMNMNIINGGAHDSGVDPVQEFMIAPIGATTFSEALRNG 185

Qy 182 AEYVHALKSVLKEGLSTGLDGGFAPSGVSTREALDLIVKAEKAGFTPGKDIALDV 241

Db 186 AEYVHALKSVLKEGLSTGLDGGFAPSGVSTREALDLIVKAEKAGFTPGKDIALDV 245

Qy 242 VASSEFFKDGTYHF--EGGQHSAAEMANVYAEVLVDAYPIVSIEDDPLQEDDWEYTNLTA 299

Db 246 PASTEFYENGKYLKGEKVYTTSEEMVTVYEEMITKYPISIEDGLDENDWGFLLTDR 305

Qy 300 IGDQVQIVGDDPFTVTPERLKEGIAKANSILVKVNOIGTTLTETFPDVMMAHRAGYTSM 359

Db 306 IGDQVQIVGDDPFTVTPERLKEGIAKANSILVKVNOIGTTLTETFPDVMMAHRAGYTAV 365

Qy 360 MSHRSGETEDTTIADLAVALNCGQIKTGAPARSDRVAKYNQLLRIBOLLGDAGVYAGRS 419

Db 366 VSHRSGETEDTTIADLAVAVNAQIKTGAPARTERVAKYNQLLRIBELLGEVAFRCLNA 425

Qy 420 F 420

Db 426 F 426

RESULT 11

F70343

enolase - Aquifex aeolicus

C;Species: Aquifex aeolicus

C;Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 16-Aug-2004

C;Accession: F70343

R;Beckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Ove

V.

Nature 392, 353-358, 1998

A;Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.

A;Reference number: A70300; MUID:98196666; PMID:9537320

A;Accession: F70343

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-426 <AQF>

A;Cross-references: UNIPROT:O66778; GB:AE000692; NID:g2983130; PIDN:AAC06738.1; PID:g2983

A;Experimental source: strain VF5

C;Genetics:

A;Gene: eno

C;Superfamily: Enolase

Query Match

61.9%; Score 1334.5; DB 2; Length 426;

RESULT 15
H75022
phosphopyruvate hydratase (EC 4.2.1.11) PAB1126 - Pyrococcus abyssi (strain Orsay)
C;Species: Pyrococcus abyssi
C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Jun-2000
C;Accession: H75022
R;anonymous, Genoscope
submitted to the EMBL Data Library, July 1999
A;Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome structure
A;Reference number: A75001
A;Accession: H75022

Search completed: September 29, 2005, 12:14:54
Job time : 29.5 secs

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2005, 23:34:02 ; Search time 88.5 Seconds
(without alignments)
2459.137 Million cell updates/sec

Title: US-10-728-947-4

Perfect score: 2155

Sequence: 1 VAIMHVFAREILDSRGNPT.....QLLGDAGVYAGRSAPFRFQG 425

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

UniProt_03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2148	99.7	425	1 ENO_CORGL	Q8nrs1 corynebacte
2	2025	94.0	425	1 ENO_COREF	Q8fqs7 corynebacte
3	1913	88.8	425	2 Q6NT61	Q8ni61 corynebacte
4	1395.5	74.0	426	1 ENO1_STRCO	Q9f2q3 streptomyc
5	1567.5	72.7	426	2 Q6ADR6	Q6adr6 leifsonia x
6	1560.5	72.4	429	2 Q741U7	Q741u7 mycobacteri
7	1553.5	72.1	429	1 ENO_MYCTU	P96377 mycobacteri
8	1553.5	72.1	429	2 Q7U0U6	Q7u0u6 mycobacteri
9	1550.5	71.9	428	2 Q82HH5	Q82hns streptomyc
10	1505.5	69.9	426	2 Q6AAB8	Q6aab8 propionibac
11	1501.5	69.7	429	1 ENO_MYCLE	Q9cd42 mycobacteri
12	1436	66.6	428	2 Q83H73	Q83h73 tropheryma
13	1436	66.6	446	2 Q83PF7	Q83ff7 tropheryma
14	1422	66.0	424	1 ENO_AGRP5	Q8ufh1 agrobacteri
15	1412.5	65.5	426	1 ENO_CAUCR	Q9a7j9 caulobacter
16	1408	65.3	429	1 ENO_THETN	Q8r967 thermoaer
17	1402	65.1	429	1 ENO_BACHD	Q9k717 bacillus ha
18	1400.5	65.0	429	2 Q84FY9	Q84fy9 methylobact
19	1398	64.9	425	1 ENO_BRUSU	Q8q0q3 brucella su
20	1397	64.8	424	1 ENO_RHIME	Q92q98 rhizobium m
21	1396	64.8	424	1 ENO_RHILLO	Q98mz3 rhizobium l
22	1391	64.5	425	1 ENO_BRUME	Q8yhf0 brucella me
23	1390.5	64.5	430	2 Q65EN2	Q65en2 bacillus li
24	1377.5	63.9	430	1 ENO_BACSU	P37869 bacillus su
25	1376.5	63.9	431	1 ENO_CLOPE	Q8xku4 clostridium
26	1375	63.8	431	2 Q815K8	Q815k8 bacillus ce
27	1373	63.7	431	2 Q81H78	Q81x78 bacillus th
28	1373	63.7	431	2 Q6HBF3	Q6hbf3 bacillus th
29	1372	63.7	432	2 Q7VDY0	Q7vdy0 prochloroco
30	1371	63.6	431	2 Q63IM2	Q63im2 bacillus ce
31	1370.5	63.6	427	1 ENO_RALSO	Q8y0b5 ralstonia s

32 1370 63.6 428 2 Q74AR6 Q74ar6 geobacter s
33 1362.5 63.2 430 2 Q7U3T1 Q7u3t1 synechococc
34 1362 63.2 431 2 Q72XV5 Q72xv5 bacillus ce
35 1361.5 63.2 430 2 Q67SV9 Q67sv9 symbiobacte
36 1359.5 63.1 430 1 ENO_LISIN P64075 listeria in
37 1359.5 63.1 430 2 Q71WX1 Q71wx1 listeria mo
38 1359.5 63.1 430 2 Q898R0 Q898r0 clostridium
39 1356.5 62.9 431 2 Q898R0 Q898r0 clostridium
40 1355.5 62.9 429 1 ENO_OCEIH O8enp5 oceanobacil
41 1355.5 62.9 434 2 Q8GDZ5 Q8gdz5 heliobacill
42 1354.5 62.9 431 1 ENO_CLOAB Q97152 clostridium
43 1350 62.6 431 2 Q7V483 Q7v483 prochloroco
44 1349 62.6 430 2 Q7V377 Q7v377 prochloroco
45 1344 62.4 422 2 Q72H85 Q72h85 thermus the

ALIGNMENTS

RESULT 1

ID_ ENO_CORGL STANDARD; PRT; 425 AA.
AC Q8NRS1;
DT 28-FEB-2003 (Rel. 41, Last Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Enolase (EC 4.2.1.11) (2-phosphoglycerate dehydratase) (2-phospho-D-
DE glycerate hydro-lyase).
GN Name=eno; OrderedLocuNames=Cgl0974, cgl111;
OS Corynebacterium glutamicum (Brevibacterium flavum).
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=1718;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
RA Nakagawa S.;
RT "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
RX PubMed=12948626; DOI=10.1016/S0168-1656(03)00154-8;
RA Kalinowski J., Bathe B., Bartels D., Bischoff N., Bott M.,
RA Burkowski A., Dusch N., Eggeling L., Eikmanns B.J., Gaigalat L.,
RA Goesmann A., Hartmann M., Huthmacher K., Kraemer R., Linke B.,
RA McHardy A.C., Meyer F., Moeckel B., Pfeifferle W., Puhler A.,
RA Rey D.A., Rueckert C., Rupp O., Sahn H., Wendisch V.F., Wiegand I.,
RA Tauch A.;
RT "The complete Corynebacterium glutamicum ATCC 13032 genome sequence
RT and its impact on the production of L-aspartate-derived amino acids
RT and vitamins.";
RL J. Biotechnol. 104:5-25(2003).
CC -I- CATALYTIC ACTIVITY: 2-phospho-D-glycerate = phosphoenolpyruvate +
CC H(2O).
CC -I- COPACTOR: Magnesium is required for catalysis and for stabilizing
CC the dimer (By similarity).
CC -I- PATHWAY: Glycolysis.
CC -I- SUBUNIT: Homodimer (By similarity).
CC -I- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -I- SIMILARITY: Belongs to the enolase family.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; AP005277; BAB98367.1; --
EMBL; BX927150; CAP19681.1; --
HSSP; P08324; 1E9I.

```

DR HAMAP; MF 00318; -, 1.
DR InterPro; IPR000941; Enolase.
DR Pfam; PF00113; Enolase_C; 1.
DR Pfam; PF03952; Enolase_N; 1.
DR PRINTS; PR00148; ENOLASE.
DR ProDom; PD000902; ENOLASE; 1.
DR TIGRFAMs; TIGR01060; eno; 1.
DR PROSITE; PS00164; ENOLASE; 1.
KW Complete proteome; Glycolysis; Lyase; Magnesium.
FT ACT_SITE 154 154 By similarity.
FT METAL 241 241 Magnesium (By similarity).
FT METAL 282 282 Magnesium (By similarity).
FT METAL 309 309 Magnesium (By similarity).
SQ SEQUENCE 425 AA; 44949 MW; B7AABD82C165FD1D CRC64;

Query Match 99.7%; Score 2148; DB 1; Length 425;
Best Local Similarity 99.5%; Pred. No. 2.5e-119;
Matches 423; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 VAEIMHVFAREILDNRGNPTVEAEVFLDDGSHGVAGVPSGASTGVHEAHELDRDGYLG 60
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1 MAEIMHVFAREILDNRGNPTVEAEVFLDDGSHGVAGVPSGASTGVHEAHELDRDGYLG 60

QY 61 KGVLKAVENNVNEEIGDELADDDORLIDEAMIKLDGTANKSRGLGANAILGVSMVAKA 120
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
61 KGVLKAVENNVNEEIGDELADDDORLIDEAMIKLDGTANKSRGLGANAILGVSMVAKA 120

QY 121 AADSAGLPLFRYIGGPNHVLVPMNMNIINGGAHDSGVVDVQEFMIAPIGAEFTFSALRN 180
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
121 AADSAGLPLFRYIGGPNHVLVPMNMNIINGGAHDSGVVDVQEFMIAPIGAEFTFSALRN 180

QY 181 GAEVYHALKSVIKEKGLSTGLDEGGFAPSGVSTREALDLIVKAIKAGFTPGKDIALAL 240
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
181 GAEVYHALKSVIKEKGLSTGLDEGGFAPSGVSTREALDLIVKAIKAGFTPGKDIALAL 240

QY 241 DVASSEFFKDGTYHFEGGQHSAAEMANVYAEVLVDAYPIVSIEDPLQDDWEGYTNLTATI 300
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
241 DVASSEFFKDGTYHFEGGQHSAAEMANVYAEVLVDAYPIVSIEDPLQDDWEGYTNLTATI 300

QY 301 GSKVQIVGDDFFVTNPERLKEGIAKKAANSILVKVNIQIGTLTETTFDAVDMHRAGYTSMW 360
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
301 GSKVQIVGDDFFVTNPERLKEGIAKKAANSILVKVNIQIGTLTETTFDAVDMHRAGYTSMW 360

QY 361 SHRSGETEDTTIADLAVALNCGOIKTGAPARSDRVAKYNOLLRIEQLLDGAGVYAGRSF 420
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
361 SHRSGETEDTTIADLAVALNCGOIKTGAPARSDRVAKYNOLLRIEQLLDGAGVYAGRSF 420

QY 421 PRFQG 425
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
421 PRFQG 425

RESULT 2
ENO COREF STANDARD; PRT; 425 AA.
AC Q8Q87;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Enolase (EC 4.2.1.11) (2-phosphoglycerate dehydratase) (2-phospho-D-
DE glycerate hydro-lyase).
GN Names: OrderedLocustNames=CE1042;
OS Corynebacterium efficiens.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
CX NCBI_TaxID=152794;
RN [1]
RP SEQUENCE FROM N. A.
RC STRAIN=Y8-314 / AJ 12310 / DSM 44549 / JCM 11189;
RX MEDLINE=22723752; PubMed=12840036; DOI=10.1101/gr.1285603;
RA Nishio Y., Nakamura Y., Kawarayashi Y., Usuda Y., Kimura E.,
RA Sugimoto S., Matsui K., Yamagishi A., Kikuchi H., Ikeo K.,
RA Gojobori T.;

```

```

RT "Comparative complete genome sequence analysis of the amino acid
RT replacements responsible for the thermostability of Corynebacterium
RL Genome Res. 13:1572-1579 (2003).
CC -!- CATALYTIC ACTIVITY: 2-phospho-D-glycerate = phosphoenolpyruvate +
CC H(2)O.
CC -!- COFACTOR: Magnesium is required for catalysis and for stabilizing
CC the dimer (By similarity).
CC -!- PATHWAY: Glycolysis.
CC -!- SUBUNIT: Homodimer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the enolase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AP005217; BAC17852.1; -.
DR HSSP; P08324; 1E91.
DR HAMAP; MF 00318; -, 1.
DR InterPro; IPR000941; Enolase.
DR Pfam; PF00113; Enolase_C; 1.
DR Pfam; PF03952; Enolase_N; 1.
DR PRINTS; PR00148; ENOLASE.
DR ProDom; PD000902; Enolase; 1.
DR TIGRFAMs; TIGR01060; eno; 1.
DR PROSITE; PS00164; ENOLASE; 1.
KW Complete proteome; Glycolysis; Lyase; Magnesium.
FT ACT_SITE 154 154 By similarity.
FT METAL 241 241 Magnesium (By similarity).
FT METAL 282 282 Magnesium (By similarity).
FT METAL 309 309 Magnesium (By similarity).
SQ SEQUENCE 425 AA; 45122 MW; 90744F1C427768C9 CRC64;

Query Match 94.0%; Score 2025; DB 1; Length 425;
Best Local Similarity 93.9%; Pred. No. 4.9e-112;
Matches 399; Conservative 9; Mismatches 17; Indels 0; Gaps 0;

QY 1 VAEIMHVFAREILDNRGNPTVEAEVFLDDGSHGVAGVPSGASTGVHEAHELDRDGYLG 60
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1 MAEIMHVFAREILDNRGNPTVEAEVFLDDGSHGVAGVPSGASTGVHEAHELDRDGYLG 60

QY 61 KGVLKAVENNVNEEIGDELADDDORLIDEAMIKLDGTANKSRGLGANAILGVSMVAKA 120
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
61 KGVLKAVENNVNEEIGDELADDDORLIDEAMIKLDGTANKSRGLGANAILGVSMVAKA 120

QY 121 AADSAGLPLFRYIGGPNHVLVPMNMNIINGGAHDSGVVDVQEFMIAPIGAEFTFSALRN 180
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
121 AADSAGLPLFRYIGGPNHVLVPMNMNIINGGAHDSGVVDVQEFMIAPIGAEFTFSALRN 180

QY 181 GAEVYHALKSVIKEKGLSTGLDEGGFAPSGVSTREALDLIVKAIKAGFTPGKDIALAL 240
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
181 GAEVYHALKSVIKEKGLSTGLDEGGFAPSGVSTREALDLIVKAIKAGFTPGKDIALAL 240

QY 241 DVASSEFFKDGTYHFEGGQHSAAEMANVYAEVLVDAYPIVSIEDPLQDDWEGYTNLTATI 300
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
241 DVASSEFFKDGTYHFEGGQHSAAEMANVYAEVLVDAYPIVSIEDPLQDDWEGYTNLTATI 300

QY 301 GSKVQIVGDDFFVTNPERLKEGIAKKAANSILVKVNIQIGTLTETTFDAVDMHRAGYTSMW 360
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
301 GSKVQIVGDDFFVTNPERLKEGIAKKAANSILVKVNIQIGTLTETTFDAVDMHRAGYTSMW 360

QY 361 SHRSGETEDTTIADLAVALNCGOIKTGAPARSDRVAKYNOLLRIEQLLDGAGVYAGRSF 420
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
361 SHRSGETEDTTIADLAVALNCGOIKTGAPARSDRVAKYNOLLRIEQLLDGAGVYAGRSF 420

QY 421 PRFQG 425
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
421 PRFQG 425

```



```

RESULT 3
Q6NI61 PRELIMINARY; PRT; 425 AA.
AC Q6NI61;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Enolase (EC 4.2.1.11).
GN Name: eno; OrderedLocusNames=DIP0917;
OS Corynebacterium diphtheriae.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=1717;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Biotype gravis / NCTC 13129;
RX MEDLINE=22955443; PubMed=14602910; DOI=10.1093/nar/gkg874;
RA Cerdeno-Tarraga A.-M., Efstathiou A., Dover L.G., Holden M.T.G.,
RA Pallen M.J., Bentley S.D., Beesa G.S., Churcher C.M., James K.D.,
RA De Zoysa A., Chillingworth T., Cronin A., Dowd L., Feltwell T.,
RA Hamlin N., Holroyd S., Jagels K., Moule S., Quail M.A.,
RA Rabinowitz E., Rutherford K.M., Thomson N.R., Unwin L.,
RA Whithead S., Barrall B.G., Parkhill J.;
RT "The complete genome sequence and analysis of Corynebacterium
diphtheriae NCTC13129."
RL Nucleic Acids Res. 31:6516-6523(2003).
CC -1- CATALYTIC ACTIVITY: 2-phospho-D-glycerate = phosphoenolpyruvate +
H(2O).
CC -1- COFACTOR: Magnesium is required for catalysis and for stabilizing
the dimer (By similarity).
CC -1- PATHWAY: Glycolysis.
CC -1- SUBUNIT: Homodimer (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the enolase family.
DR EMBL; BX248356; CAB9433.1; -.
DR GO; GO:0000015; C:phosphopyruvate hydratase complex; IEA.
DR GO; GO:0016829; F:lyase activity; IEA.
DR GO; GO:0000287; F:magnesium ion binding; IEA.
DR GO; GO:0004634; F:phosphopyruvate hydratase activity; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR000941; Enolase.
DR Pfam; PF00113; Enolase C; 1.
DR Pfam; PF03952; Enolase N; 1.
DR PRINTS; PR00148; ENOLASE.
DR ProDom; PD000902; Enolase; 1.
DR TIGRFAMs; TIGR01060; eno; 1.
DR PROSITE; PS00164; ENOLASE; 1.
KW Complete proteome; Glycolysis; Lyase; Magnesium.
SQ SEQUENCE 425 AA; 45029 MW; 85BF477BE7ECDCA6 CRC64;

Query Match 88.8%; Score 1913; DB 2; Length 425;
Best Local Similarity 87.8%; Pred. No. 2.1e-105;
Matches 373; Conservative 23; Mismatches 29; Indels 0; Gaps 0;

QY 1 VAEIMVFAREILLDSRNPVTEAEVFLDDGSHGVAGVPSCAGTGVVHAHELRRGGRYLG 60
Db 1 MADIMHVFAREILLDSRNPVTEAEVFLDDGSHGVAGVPSCAGTGVVHAHELRRGGRYLG 60
QY 61 KGVLKAVENYNEIGDELAGEADDQRLIDEAMIKLDTGTANKSRLGANAILGVSMVAKA 120
Db 61 KGVLNANVNNVEEIAAIAEADDQRLIDQAMIALDGTENKSRNLGANAILGVSIYAKA 120
QY 121 AADSAGLPLPRYTGPNNAHVLVPMNMINGGAHADSGVDVQEFMIAPICAEFTFSBALRN 180
Db 121 AAEASAGLPLRYTGGPNNAHVLVPMNMVINGGAHADSGVDVQEFMIAPICAEFTFSBALRM 180
QY 181 GAEVYHALKSVIEKGLSTGLGDEGGPAPSVGSTREALDLIVKAEKAGTTPCKDTALAL 240
Db 181 GAEVYHSLKSVIKSLGSLTGLGDEGGPAPSVGSTREALDLIVKAEKAGTTPCKDTALAL 240
QY 241 DVASSBFFKDGTYHFEGGQHSAAEMANVYAEVDAVPVIVSIEDPLQEDDDWEGYTALTAI 300

```

U

```

RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
CC -I- CATALYTIC ACTIVITY: 2-phospho-D-glycerate = phosphoenolpyruvate +
CC H(2)O.
CC
CC -I- COFACTOR: Magnesium is required for catalysis and for stabilizing
CC the dimer (By similarity).
CC -I- PATHWAY: Glycolysis.
CC -I- SUBUNIT: Homodimer (By similarity).
CC -I- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -I- SIMILARITY: Belongs to the enolase family.
DR EMBL: AE017230; AA033307.1; -.
DR GO: GO:0000015; C:phosphopyruvate hydratase complex; IEA.
DR GO: GO:0004634; F:phosphopyruvate hydratase activity; IEA.
DR GO: GO:0006096; P:glycolysis; IEA.
DR InterPro: IPR000941; Enolase.
DR Pfam: PF001113; Enolase_C; 1.
DR Pfam: PF03952; Enolase_N; 1.
DR PRINTS: PR00148; ENOLASE.
DR ProDom: PD000902; Enolase; 1.
DR TIGRFAMs: TIGR01060; eno; 1.
DR PROSITE: PS00164; ENOLASE; 1.
KW Complete proteome; Glycolysis; Lyase; Magnesium.
SQ SEQUENCE 429 AA; 44873 MW; C5B7C0460F01D703 CRC64;

Query Match 72.4%; Score 1560.5; DB 2; Length 429;
Best Local Similarity 72.9%; Pred. No. 1.6e-84;
Matches 307; Conservative 47; Mismatches 66; Indels 1; Gaps 1;

QY 4 IMHVFAREILDSRGNTPTVEAEVFLDDGSHGVAGVPSCASTGVHEAHLRGGDRYLKGV 63
DB 4 IEQVGAIREILDSRGNTPTVEAEIALDTGTFARAAVPSGASTGEHEAHLRGGDRYGGKV 63

QY 64 LKAIVENNERIGDELAGEADDDORLIDEAMIKLDGTANKSRGLGANALIGVSMVAKAAAD 123
DB 64 QKAVQAVLDEIGPANGVGLNADDQRLVDQALVDLDTGTPDKSRGLGNALIGVSLAVAKAAD 123

QY 124 SAGLPLFRYIGGNNAHVLPPVMNIIINGGAHSGVDVQBFMIAPIGAEFTSBAALRWAE 183
DB 124 SABLPLFRYIGGNNAHLPVPMNIIINGGAHADTAVDIQBFMVAPIGAFSAELRWAE 183

QY 184 VYHALKSIVKEGLSTGLGDEGGFAPSGVSTREALDLIVKAIKAGTPGKDTALALDVA 243
DB 184 VYHSLSVLKKGLSTGLGDEGGFADPVAGTTAALDLIGRAIESAGFKLGTVALALDAA 243

QY 244 SSEFFKDGDT-YHPEGGGSHSAEAMVAVAEIVDAPYIVSIEDPLEQDDWEGYTNLTATIGD 302
DB 244 ATFFYSDGTGYKEGSTRTAQMAEFYAGLLGAYPLVSIEDPLESDDDWDGWAALTASIGD 303

QY 303 KVOIVGDDFFVTNPERLKEGIKAANSILVKVQNQIGTLTETFDVDMHRAGYTSMMSH 362
DB 304 RVQLVGDDVFTNPERLEEGIEKGVANALLVKVQNQIGTLTETLDAVALAHSGVRYTMSH 363

QY 363 RSGETEDTTIADLAVLNCCQIKTGAPARSDRVAKYNQLRIEQLIGDAGVYAGRSAPFR 422
DB 364 RSGETEDTTIADLAVAGVSGQIKTGAPARSDRVAKYNQLRIEALGDAARYAGDLAPFR 423

QY 423 F 423
DB 424 F 424

RESULT 7
ID ENO_MYCTU STANDARD; PRT; 429 AA.
AC P96377;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 25-JAN-2005 (Rel. 46, Last annotation update)
DE Enolase (EC 4.2.1.11) (2-phosphoglycerate dehydratase) (2-phospho-D-
DE glycerate hydro-lyase).
DN Name=eno; OrderedLocuNames=Rv1023, MT1051; ORFNames=MTCY10G2.26c;
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteriales; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.

```

```

OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37Rv;
RX MEDLINE=98295987; PubMed=9634230; DOI=10.1038/31159;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C.M.,
RA Harris D.E., Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III,
RA Tekala F., Badcock K., Basham D., Brown D., Chillingworth T.,
RA Connor R., Davies R.M., Devlin K., Feltwell T., Gentles S., Hamlin N.,
RA Holroyd S., Hornsby T., Jagels K., Krogh A., McLean J., Moule S.,
RA Murphy L.D., Oliver S., Osborne J., Quail M.A., Rajandream M.A.,
RA Rogers J., Rutter S., Seeger K., Skelton S., Squares R.,
RA Stoughton J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 393:537-544 (1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RX MEDLINE=22206494; PubMed=12218036;
RX DOI=10.1128/JB.184.19.5479-5490.2002;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J.D., DeBoy R.T., Dodson R.J., Gwin M.L., Haft D.H.,
RA Hickey E.K., Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D.,
RA Salzberg S.L., Delcher A., Utterback T.R., Weidman J.F., Khouri H.M.,
RA Gill J., Mikula A., Bishai W., Jacobs W.R. Jr., Venter J.C.,
RA Fraser C.M.;
RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RL J. Bacteriol. 184:5479-5490 (2002).
CC -I- CATALYTIC ACTIVITY: 2-phospho-D-glycerate = phosphoenolpyruvate +
CC H(2)O.
CC -I- COFACTOR: Magnesium is required for catalysis and for stabilizing
CC the dimer (By similarity).
CC -I- PATHWAY: Glycolysis.
CC -I- SUBUNIT: Homodimer (By similarity).
CC -I- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -I- SIMILARITY: Belongs to the enolase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: BX842575; CAB06856.1; -.
CC EMBL: AE000516; AAK45302.1; -.
CC PIR: B70623; B70623.
CC HSSP: O9NDH8; 1OEP.
CC TIGR: MT1051; -.
CC Tuberculist; Rv1023; -.
CC HAMAP; MF_00318; -.
CC InterPro: IPR000941; Enolase.
CC Pfam: PF00113; Enolase_C; 1.
CC Pfam: PF03952; Enolase_N; 1.
CC PRINTS: PR00148; ENOLASE.
CC ProDom: PD000902; Enolase; 1.
CC TIGRFAMs: TIGR01060; eno; 1.
CC PROSITE: PS00164; ENOLASE; 1.
KW Complete proteome; Glycolysis; Lyase; Magnesium.
FT ACT_SITE 154 154 By similarity.
FT METAL 241 241 Magnesium (By similarity).
FT METAL 283 283 Magnesium (By similarity).
FT METAL 310 310 Magnesium (By similarity).
SQ SEQUENCE 429 AA; 44961 MW; 5AC59531AC9F00E8 CRC64;

Query Match 72.1%; Score 1553.5; DB 1; Length 429;
Best Local Similarity 73.2%; Pred. No. 4.3e-84;
Matches 308; Conservative 42; Mismatches 70; Indels 1; Gaps 1;

QY 4 IMHVFAREILDSRGNTPTVEAEVFLDDGSHGVAGVPSCASTGVHEAHLRGGDRYLKGV 63

```

Db 4 IEQVAREILDSRGNTVEVEVALIDGTFARAAVPSGASTGBHEAVEFLRDGGDRYGGKGV 63
 QY 64 LKAVERNBEIGDELGLAGLEADQRLIDEAMIKLDGTANKSRGCAILGVSMVAKAAAD 123
 Db 64 QKAVQAVLDEIGPAVIGLGNADQRLVDLQALVDLDTGTPDKSRGLGNAILGVSLAVAKAAAD 123
 QY 124 SAGLPLFRYIGGPNHVLVPMNNIINGGAHADVQVDFVQEFMIAPIGAEFTSEALRWGAE 183
 Db 124 SAGLPLFRYIGGPNHVLVPMNNIINGGAHADVQVDFVQEFMIAPIGAEFTSEALRWGAE 183
 QY 184 VYHALKSVLKEGLSTGLGDEGFAPSVGSTREALDLIVKATEKAGFTPGKDIALALDVA 243
 Db 184 VYHALKSVLKEGLSTGLGDEGFAPSVGSTREALDLIVKATEKAGFTPGKDIALALDVA 243
 QY 244 SFEFFKDG-T-YHFEQGQSHAAEMANVYAEVLVDAYPIVSTEDPLQDDWEGYTNLTATIGD 302
 Db 244 ATEFFTDGTGYVFEFTTADQMTFYGALLGAYPLVSTEDPLSDDWDGWAALTASIGD 303
 QY 303 KVOIVGDDFFVTNPERLKEGIAKKAANSILVKVNOIGTLTETFDVDMHAHAGYTSMMSH 362
 Db 304 RVOIVGDDIFVTNPERLEEGIERGVANALLVKVNOIGTLTETDLAVTLAHHGGYRTMISH 363
 QY 363 RSGETEDTTIADLAVALNCGQIKTGAPARSDRVAKYNOLLRIEQLLDGAGVYAGRSAPPR 422
 Db 364 RSGETEDTTIADLAVAIAGSGQIKTGAPARSDRVAKYNOLLRIEALGDAARYAGDLAFPR 423
 QY 423 F 423
 Db 424 F 424

RESULT 8

Q7U006
 ID Q7U006 PRELIMINARY; PRT; 429 AA.
 AC Q7U006
 DT 01-OCT-2003 (TremBLrel. 25, Created)
 DT 01-OCT-2003 (TremBLrel. 25, Last sequence update)
 DT 01-MAR-2004 (TremBLrel. 26, Last annotation update)
 DE PROBABLE ENOLASE ENO (EC 4.2.1.11).
 GN Name=eno; OrderedLocusNames=Mb1051;
 OS Mycobacterium bovis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1765;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AF2122/97;
 RX MEDLINE=22709107; PubMed=12788972; DOI=10.1073/pnas.1130426100;
 RA Garnier T., Eigmeier K., Camus J.-C., Medina N., Mansoor H.,
 RA Pryor M., Duthoy S., Grondin S., Lacroix C., Monsemp C., Simon S.,
 RA Harris B., Actin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
 RA Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
 RT "The complete genome sequence of Mycobacterium bovis."
 RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).
 CC -1- CATALYTIC ACTIVITY: 2-phospho-D-glycerate = phosphoenolpyruvate +
 H(2)O.
 CC -1- COFACTOR: Magnesium is required for catalysis and for stabilizing
 CC the dimer (By similarity).
 CC -1- PATHWAY: Glycolysis.
 CC -1- SUBUNIT: Homodimer (By similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: Belongs to the enolase family.
 DR EMBL: BX248337; CAD93911.1; -.
 DR HSP: P08324; I9E1.
 DR GO: GO:0000015; C:phosphopyruvate hydratase complex; IEA.
 DR GO: GO:0016829; F:lyase activity; IEA.
 DR GO: GO:0000287; F:magnesium ion binding; IEA.
 DR GO: GO:0004634; F:phosphopyruvate hydratase activity; IEA.
 DR GO: GO:0006096; P:glycolysis; IEA.
 DR InterPro: IPR000941; Enolase.
 DR Pfam: PF00113; Enolase_C; 1.
 DR Pfam: PF03952; Enolase_N; 1.

DR ProDom: PD000902; Enolase; 1.
 DR TIGRFAMs: TIGR01060; eno; 1.
 DR PROSITE: PS00164; ENOLASE; 1.
 KW Complete proteome; Glycolysis; Lyase; Magnesium.
 SQ SEQUENCE 429 AA; 44862 MW; F808FB8E0D891C9A CRC64;

Query Match 72.1%; Score 1553.5; DB 2; Length 429;
 Best Local Similarity 73.2%; Pred. No. 4.3e-84;
 Matches 308; Conservative 42; Mismatches 70; Indels 1; Gaps 1;

QY 4 IMHVAREILDSRGNTVEVEVALIDGTFARAAVPSGASTGBHEAVEFLRDGGDRYGGKGV 63
 Db 4 IEQVAREILDSRGNTVEVEVALIDGTFARAAVPSGASTGBHEAVEFLRDGGDRYGGKGV 63
 QY 64 LKAVERNBEIGDELGLAGLEADQRLIDEAMIKLDGTANKSRGCAILGVSMVAKAAAD 123
 Db 64 QKAVQAVLDEIGPAVIGLGNADQRLVDLQALVDLDTGTPDKSRGLGNAILGVSLAVAKAAAD 123
 QY 124 SAGLPLFRYIGGPNHVLVPMNNIINGGAHADVQVDFVQEFMIAPIGAEFTSEALRWGAE 183
 Db 124 SAGLPLFRYIGGPNHVLVPMNNIINGGAHADVQVDFVQEFMIAPIGAEFTSEALRWGAE 183
 QY 184 VYHALKSVLKEGLSTGLGDEGFAPSVGSTREALDLIVKATEKAGFTPGKDIALALDVA 243
 Db 184 VYHALKSVLKEGLSTGLGDEGFAPSVGSTREALDLIVKATEKAGFTPGKDIALALDVA 243
 QY 244 SFEFFKDG-T-YHFEQGQSHAAEMANVYAEVLVDAYPIVSTEDPLQDDWEGYTNLTATIGD 302
 Db 244 ATEFFTDGTGYVFEFTTADQMTFYGALLGAYPLVSTEDPLSDDWDGWAALTASIGD 303
 QY 303 KVOIVGDDFFVTNPERLKEGIAKKAANSILVKVNOIGTLTETFDVDMHAHAGYTSMMSH 362
 Db 304 RVOIVGDDIFVTNPERLEEGIERGVANALLVKVNOIGTLTETDLAVTLAHHGGYRTMISH 363
 QY 363 RSGETEDTTIADLAVALNCGQIKTGAPARSDRVAKYNOLLRIEQLLDGAGVYAGRSAPPR 422
 Db 364 RSGETEDTTIADLAVAIAGSGQIKTGAPARSDRVAKYNOLLRIEALGDAARYAGDLAFPR 423
 QY 423 F 423
 Db 424 F 424

RESULT 9

Q82HH5
 ID Q82HH5 PRELIMINARY; PRT; 428 AA.
 AC Q82HH5
 DT 01-JUN-2003 (TremBLrel. 24, Created)
 DT 01-JUN-2003 (TremBLrel. 24, Last sequence update)
 DT 01-MAR-2004 (TremBLrel. 26, Last annotation update)
 DE Putative enolase.
 GN Name=eno; OrderedLocusNames=SAV3533;
 OS Streptomyces avermitilis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Streptomycineae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=33903;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MA-4680;
 RX MEDLINE=21477403; PubMed=11572948; DOI=10.1073/pnas.211433198;
 RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
 RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Oonoe T.,
 RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
 RT "Genome sequence of an industrial microorganism Streptomyces
 RT avermitilis: deducing the ability of producing secondary
 RT metabolites";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MA-4680;
 RX MEDLINE=22608306; PubMed=12692562;
 RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
 RA Sakaki Y., Hattori M., Omura S.;

"Complete genome sequence and comparative analysis of the industrial microorganism *Streptomyces avermitilis*."; Nat. Biotechnol. 21:526-531(2003).
 CC -1- CATALYTIC ACTIVITY: 2-phospho-D-glycerate = phosphoenolpyruvate + H₂O.
 CC -1- COFACTOR: Magnesium is required for catalysis and for stabilizing the dimer (By similarity).
 CC -1- PATHWAY: Glycolysis.
 CC -1- SUBUNIT: Homodimer (By similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: Belongs to the enolase family.
 DR HMBL; AP005035; BAC71245.1; -.
 DR HSP; Q9NDH8; IOEP.
 DR GO; GO:0000015; C:phosphopyruvate hydratase complex; IEA.
 DR GO; GO:0016829; F:lyase activity; IEA.
 DR GO; GO:000287; F:magnesium ion binding; IEA.
 DR GO; GO:0004634; F:phosphopyruvate hydratase activity; IEA.
 DR GO; GO:0006096; P:glycolysis; IEA.
 DR InterPro; IPR000941; Enolase.
 DR Pfam; PF01113; Enolase_C; 1.
 DR PRINTS; PR00148; ENOLASE.
 DR ProDom; PD000902; Enolase; 1.
 DR TIGRFAMs; TIGR01060; eno; 1.
 DR PROSITE; PS00164; ENOLASE; 1.
 KW Complete proteome; Glycolysis; Lyase; Magnesium.
 SQ SEQUENCE 428 AA; 45869 MW; 095A3976F6C5F70D CRC64;
 Query Match 71.9%; Score 1550.5; DB 2; Length 428;
 Best Local Similarity 73.5%; Pred. No. 6.4e-84;
 Matches 308; Conservative 43; Mismatches 67; Indels 1; Gaps 1;
 QY 7 VFAREILDSRGNTVEAEVFLDDGSHGVAGVPSCASTGVHEAHLRGG-DRYLKGVLK 65
 DB 7 VFAREILDSRGNTVEAEVFLDDGSHGVAGVPSCASTGVHEAHLRGG-DRYLKGVLK 66
 QY 66 AVENVNEEIGDELADQRLIDEAMIKLDGTANKSRGLGANAILGVSMVAKAAADSA 125
 DB 67 AVLAVIEQIGPELVGYDATEQRLIDQAMFDDATDNKSLGANAILGVSLAVAHAAEAS 126
 QY 126 GLPLFRYIGGPNHVLVPMNMIINGGAHDSGVVDVQEFMIAPIGASTFSEALRNGAEVY 185
 DB 127 DLPLFRYLGPNHVLVPMNMIINGGSHADSNVDIQEFMIAPIGASTFSEALRNGAEVY 186
 QY 186 HALKSVIKKGLSTGLGDEGFAPSVGSTRALDLIVKAIKAGFTPGKDIALADVASS 245
 DB 187 HTLUKKVLKTLGSLGDEGFAPNLESNRAALDLIIKQAGYIPGEQIALADVASS 246
 QY 246 EFPKDGTYHFGGQHSAAEMANVYAEVLVDAYPIVSIEDPLEDDWEGYTNLTATIGDKVQ 305
 DB 247 EFKDGKYEFGKRSAAEMTEYEEVLVSAYPLVSIEDPLEDDWAGKVIITDKLGDVQ 306
 QY 306 IVGDDPFVTNPERLKEGIAKKAANSILVKVNIQIGTLTFTDAMDMAHAGYTSNMHSRSG 365
 DB 307 IVGDDLVFTNPERLARGIEGSANALLVKVNIQIGSLTETLDVELAQNRGFKCMHSRSG 366
 QY 366 ETEDTTIADLAVNLCCQIKTGAPSRDRVAKYNOLLRIEQLLDGAGVYAGRAFPFRQ 424
 DB 367 ETEDVTIADLAVAVNCQIKTGAPSRDRVAKYNOLLRIEQLLDGAGVYAGRAFPFR 425
 RESULT 10
 Q6AAB8 PRELIMINARY; PRT; 426 AA.
 AC Q6AAB8
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
 DE Enolase (EC 4.2.1.11)
 GN OrderedLocusNames=PPA0545;
 OS Propionibacterium acnes.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Propionibacterineae; Propionibacteriaceae; Propionibacterium.

NCBI_TaxID=1747;
 [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=KPA171202 / DSM 16379;
 RX PubMed=15286373; DOI=10.1126/science.1100330;
 RA Brueggemann H., Henne A., Hoster F., Liesegang H., Wierze A.,
 RA Strittmatter A., Hujer S., Duerre P., Gottschalk G.;
 RT "The complete genome sequence of *Propionibacterium acnes*, a commensal of human skin";
 RL Science 305:671-673(2004).
 CC -1- CATALYTIC ACTIVITY: 2-phospho-D-glycerate = phosphoenolpyruvate + H₂O.
 CC -1- COFACTOR: Magnesium is required for catalysis and for stabilizing the dimer (By similarity).
 CC -1- PATHWAY: Glycolysis.
 CC -1- SUBUNIT: Homodimer (By similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: Belongs to the enolase family.
 DR EMBL; AE017283; AAT82298.1; -.
 DR GO; GO:0000015; C:phosphopyruvate hydratase complex; IEA.
 DR GO; GO:0016829; F:lyase activity; IEA.
 DR GO; GO:0004634; F:phosphopyruvate hydratase activity; IEA.
 DR GO; GO:0006096; P:glycolysis; IEA.
 DR InterPro; IPR000941; Enolase.
 DR Pfam; PF01113; Enolase_C; 1.
 DR PRINTS; PR00148; ENOLASE.
 DR ProDom; PD000902; Enolase; 1.
 DR TIGRFAMs; TIGR01060; eno; 1.
 DR PROSITE; PS00164; ENOLASE; 1.
 KW Complete proteome; Glycolysis; Lyase; Magnesium.
 SQ SEQUENCE 426 AA; 45530 MW; 39FD0E9A2941123F CRC64;
 Query Match 69.9%; Score 1505.5; DB 2; Length 426;
 Best Local Similarity 68.9%; Pred. No. 2.9e-81;
 Matches 293; Conservative 59; Mismatches 73; Indels 1; Gaps 1;
 QY 1 VAEIMHVFAREILDSRGNTVEAEVFLDDGSHGVAGVPSCASTGVHEAHLRGG-DRYL 59
 DB 1 MATIEFIEAREILDSRGNTVEAEVFLDDGSHGVAGVPSCASTGVHEAHLRGG-DRYL 60
 QY 60 KGKVLKAVENVNEEIGDELADQRLIDEAMIKLDGTANKSRGLGANAILGVSMVAK 119
 DB 61 KGKVLKAVENVNEEIGDELADQRLIDEAMIKLDGTANKSRGLGANAILGVSMVAK 120
 QY 120 AAADSGALPLFRYIGGPNHVLVPMNMIINGGAHDSGVVDVQEFMIAPIGASTFSEALR 179
 DB 121 AAADCAELPLYQYLGPNHVLVPMNMIINGGAHDSGVVDVQEFMIAPIGASTFSEALR 180
 QY 180 NGAENVHAKSVIKKGLSTGLGDEGFAPSVGSTRALDLIVKAIKAGFTPGKDIALA 239
 DB 181 WGAENVHAKSVIKKGLSTGLGDEGFAPSVGSTRALDLIVKAIKAGFTPGKDIALA 240
 QY 240 LDVASSEFFKDGTYHFGGQHSAAEMANVYAEVLVDAYPIVSIEDPLEDDWEGYTNLTAT 299
 DB 241 LDVASSEFFKDGTYHFGGQHSAAEMANVYAEVLVDAYPIVSIEDPLEDDWEGYTNLTAT 300
 QY 300 IGDKVQIVGDDPFVTNPERLKEGIAKKAANSILVKVNIQIGTLTFTDAMDMAHAGYTSM 359
 DB 301 LGSEKIQIVGDDLVFTNPERLARGIEGSANALLVKVNIQIGSLTETLDVELAHRNGYRCM 360
 QY 360 MSHRSGETEDTTIADLAVNLCCQIKTGAPSRDRVAKYNOLLRIEQLLDGAGVYAGRA 419
 DB 361 MSHRSGETEDTTIADLAVNLCCQIKTGAPSRDRVAKYNOLLRIEQLLDGAGVYAGRA 420
 QY 420 FPRFQ 424
 DB 421 FPRFQ 425
 RESULT 11
 ENO_MYCLE
 ID ENO_MYCLE STANDARD; PRT; 429 AA.

Q9CD42;	124	SAGLPLFRYIGGPNNAHLVLPVPMNIINGGAHAGSDVDVQEFMIAPIGAEFSEALRNGAE	183
DT 28-FEB-2003 (Rel. 41, Created)			
DT 28-FEB-2003 (Rel. 41, Last sequence update)			
DT 25-OCT-2004 (Rel. 45, Last annotation update)			
DE Enolase (EC 4.2.1.11) (2-phosphoglycerate dehydratase) (2-phospho-D-glycerate hydro-lyase).	124	SAELPLFRYIGGNAHLVLPVPMNINLNGGAHADTAVDVQEFMVAPIGAPSFVEALRWGAE	183
GN Name=eno; OrderedLocusNames=ML0255;			
OS Mycobacterium leprae.	184	VYHALKSIVKEKGLSTGLGDEGFPASVSGSTRALDLIVKAIKAGTTPGKDIALALDVA	243
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;			
OC Corynebacteriineae; Mycobacteriaceae; Mycobacterium.	184	VYHALKSIVLKKGLSTGLGDEGFPASVAGTAAALDLVTLAIKAGTTPGKDIALALDVA	243
OX NCBI_TaxID=1769;			
RN [1]			
RP SEQUENCE FROM N.A.			
RC STRAIN=TN;			
RX MEDLINE=21128732; PubMed=11234002; DOI=10.1038/35059006;	244	SSEFFKDGDT-YHFEGGGSHAAEMANVYAEIVSVISIEDPLQEDDWEQVNTLTATIGD	302
RA Cole S.T., Eglmeier K., Parkhill J., James K.D., Thomson N.R.,			
RA Wheeler P.R., Honore N., Garnier T., Churcher C.M., Harris D.E.,			
RA Mungall K.L., Basham D., Brown D., Chillingworth T., Connor R.,	244	ATEFYTDGIGYHEGTHADQMTQTEFYADLLGSGYPLVISEDPLSEDDWGWAALTASIGE	303
RA Davies R.M., Hohnsby T., Duthoy S., Feltwell T., Fraser A., Hamlin N.,			
RA Holroyd S., Horneby T., Jagels K., Lacroix C., Maclean J., Moule S.,			
RA Murphy L.D., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,			
RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,			
RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,			
RA Barrell B.G.;			
RT "Massive gene decay in the leprosy bacillus.";			
RL Nature 409:1007-1011(2001).			
CC -1- CATALYTIC ACTIVITY: 2-phospho-D-glycerate = phosphoenolpyruvate + H(2)O.			
CC -1- COPACTOR: Magnesium is required for catalysis and for stabilizing the dimer (By similarity).			
CC -1- PATHWAY: Glycolysis.			
CC -1- SUBUNIT: Homodimer (By similarity).			
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).			
CC -1- SIMILARITY: Belongs to the enolase family.			
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC -----			
CC EMBL; AL583917; CAC29763.1; ALT_INIT.			
CC HSP; Q9NDH8; 10EP.			
CC Leprosa; ML0255; -.			
CC HAMAP; MF_00318; -; 1.			
CC InterPro; IPR000941; Enolase.			
CC Pfam; PF00113; Enolase_C; 1.			
CC Pfam; PF03952; Enolase_N; 1.			
CC PRINTS; PR00148; ENOLASE.			
CC ProDom; PD000902; Enolase; 1.			
CC TIGRFAMs; TIGR01060; eno; 1.			
CC PROSITE; PS00164; ENOLASE; 1.			
CC Complete proteome; Glycolysis; Lyase; Magnesium.			
FT ACT_SITE 154 154 By similarity.			
FT METAL 241 241 Magnesium (By similarity).			
FT METAL 283 283 Magnesium (By similarity).			
FT METAL 310 310 Magnesium (By similarity).			
SQ SEQUENCE 429 AA; 45201 MW; A8299C1EA44617B5 CRC64;			
Query Match 59.7%; Score 1501.5; DB 1; Length 429;			
Best Local Similarity 71.0%; Pred. No. 5.1e-81;			
Matches 299; Conservative 47; Mismatches 74; Indels 1; Gaps 1;			
QY 4 IMHVFAREILDSRGNTVPEAEVFLDDSGHGVAGVPSGAGTGVHEAELRDGGDYLKGV 63			
Db 4 IEQVGAREILDSRGNTVPEAEVFLDDSGHGVAGVPSGAGTGVHEAELRDGGDYLKGV 63			
QY 64 LKAVERNNEIGDELGLAGLADDQRLIDEAMIKLDGTANKSRGANAILGVSAVAKAAAD 123			
Db 64 KRAVDVAVLDEIGPVWIGLNDQRLIDQELLDGTPDKSRGLGNAILGVSLAVAKAAAD 123			

QY	124	SAGLPLFRYIGGPNNAHLVLPVPMNIINGGAHAGSDVDVQEFMIAPIGAEFSEALRNGAE	183
Db	124	SAELPLFRYIGGNAHLVLPVPMNINLNGGAHADTAVDVQEFMVAPIGAPSFVEALRWGAE	183
QY	184	VYHALKSIVKEKGLSTGLGDEGFPASVSGSTRALDLIVKAIKAGTTPGKDIALALDVA	243
Db	184	VYHALKSIVLKKGLSTGLGDEGFPASVAGTAAALDLVTLAIKAGTTPGKDIALALDVA	243
QY	244	SSEFFKDGDT-YHFEGGGSHAAEMANVYAEIVSVISIEDPLQEDDWEQVNTLTATIGD	302
Db	244	ATEFYTDGIGYHEGTHADQMTQTEFYADLLGSGYPLVISEDPLSEDDWGWAALTASIGE	303
QY	303	KVQIVGDDFFVTPNPERLKEGIAKKAANSILVKVNIQIGTLTETPDVDMAHRAGYTSMMSH	362
Db	304	QVQIVGDDIFATNPERLEEGIRGVANALLVKVNIQIGTLTETLEAVALAHSQYRTMISH	363
QY	363	RSGETEDTTIADLAVALNCQIKTGAPARSDRVAKYNOLLRIEQLLDGAGVYAGRAFPR	422
Db	364	RSGETEDTMIADLVVALGSGQIKTGAPARSKVAKYNOLLRIEQLLDGAGVYAGRAFPR	423
QY	423	F 423	
Db	424	Y 424	
RESULT 12			
Q83H73			
AC	Q83H73	PRELIMINARY; PRT; 428 AA.	
DT	01-JUN-2003 (TrEMBLrel. 24, Created)		
DT	01-JUN-2003 (TrEMBLrel. 24, Last sequence update)		
DT	01-MAR-2004 (TrEMBLrel. 26, Last annotation update)		
DE	Enolase (EC 4.2.1.11).		
GN	Name=eno; OrderedLocusNames=TW793;		
OS	Tropheryma whipplei (strain TW08/27) (Whipple's bacillus).		
OC	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;		
OC	Micrococcineae; Cellulomonadaceae; Tropheryma.		
OX	NCBI_TaxID=218496;		
RN	[1]_TaxID=218496;		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=TW08/27;		
RX	MEDLINE=22495039; PubMed=12606174; DOI=10.1016/S0140-6736(03)12597-4;		
RA	Bentley S.D., Maiwald M., Murphy L.D., Pallen M.J., Yeats C.A.,		
RA	Dover L.G., Norbertczak H.T., Besra G.S., Quail M.A., Harris D.E.,		
RA	von Herbay A., Goble A., Rutter S., Squares R., Squares S.,		
RA	Barrell B.G., Parkhill J., Rellman D.A.;		
RT	"Sequencing and analysis of the genome of the Whipple's disease bacterium Tropheryma whipplei.";		
RL	Lancet 361:637-644(2003).		
CC	-1- CATALYTIC ACTIVITY: 2-phospho-D-glycerate = phosphoenolpyruvate + H(2)O.		
CC	-1- COPACTOR: Magnesium is required for catalysis and for stabilizing the dimer (By similarity).		
CC	-1- PATHWAY: Glycolysis.		
CC	-1- SUBUNIT: Homodimer (By similarity).		
CC	-1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).		
CC	-1- SIMILARITY: Belongs to the enolase family.		
DR	EMBL; BX251412; CAD67452.1; -.		
DR	HSP; Q9NDH8; 10EP.		
DR	GO; GO:0000015; C:phosphoryruvate hydratase complex; IEA.		
DR	GO; GO:0016829; F:lyase activity; IEA.		
DR	GO; GO:0000287; F:magnesium ion binding; IEA.		
DR	GO; GO:0004634; F:phosphoryruvate hydratase activity; IEA.		
DR	GO; GO:0006096; P:glycolysis; IEA.		
DR	InterPro; IPR000941; Enolase.		
DR	Pfam; PF00113; Enolase_C; 1.		
DR	Pfam; PF03952; Enolase_N; 1.		
DR	PRINTS; PR00148; ENOLASE.		
DR	ProDom; PD000902; Enolase; 1.		
DR	TIGRFAMs; TIGR01060; eno; 1.		
DR	PROSITE; PS00164; ENOLASE; 1.		
DR	Complete proteome; Glycolysis; Lyase; Magnesium.		
FT	ACT_SITE 154 154 By similarity.		
FT	METAL 241 241 Magnesium (By similarity).		
FT	METAL 283 283 Magnesium (By similarity).		
FT	METAL 310 310 Magnesium (By similarity).		
SQ	SEQUENCE 428 AA; 45961 MW; 356D63DCCC24B581 CRC64;		

```

Query Match      66.6%; Score 1436; DB 2; Length 428;
Best Local Similarity 67.9%; Pred. No. 3.9e-77;
Matches 286; Conservative 51; Mismatches 82; Indels 2; Gaps 2;

QY 7 VFAREILDSRGNPTVEAEVFLDDGSHGVAGVPSGASTGVHAEHLRD-GGDRLYLGKGVLK 65
Db 7 VHARQILDSRGTPTEVEVTELDGSMGRSAVPSGASTGAFENAEHLRDQDNNEVILGKGVTR 66
QY 66 AVENVNEEIGDELAGLEADDQRLIDEAMIKLDGTANKSRILGANAILGVSMVAKAAADSA 125
Db 67 AVRSVNSEIAPVLIGFDQVGLDHRMIELDGTNNKSRILGANAILGVSLASAAARAA 126
QY 126 GLPLFRYIGGNPAHVLFPVPMNIIINGGAHDSGVVDVQEFMIAPIGAEFTFSEALRNGAEVY 185
Db 127 DLSLFRYLGSPSSRILPVPMMNIIINGGAHDTGVDIOEFMILPVGARFSFSESLRWGCEVY 186
QY 186 HALKSUVKEKGLSTGLDGGFPAPSGVSTREALDLIVKAIEKAGFTPGKDIALLDVASS 245
Db 187 HSLKVQLRESGLSGLDGGFPALRSNRALTDLILSAIEKAGFSPGIDIVLALDIAAS 246
QY 246 EFKE-DGTYHFEFGQHSAAEMANVYAEIVDAYPTVSTEDPLOEDDWEYTNLTATIGDKV 304
Db 247 EFCAPGHYRFEKDTSDLSIYEGLLSSYPLVSTIEDPLDQDDWEGYRTLTTHIGDRV 306
QY 305 QIVGDDDFVTNPNRLKEGIAKKAANSILVKVNOIGTLTETFDVADMAHRAAGYTSMMSHRS 364
Db 307 QIVGDDDLFVNTSRLSRGSIQSGVANSILVKVNOIGTLTETFDVADMAAKGGYTAVALSHRS 366
QY 365 GETEDTTIADLAVALNCQIKTGAPARSDRVAKYNQLLRLEQLLGDAGVYAGRSAPFRFQ 424
Db 367 GETEDTTIADMAVATNCQIKTGAPARGERIAKYNQLLRLEELKGRSARYAGWLSYPRWQ 426
QY 425 G 425
Db 427 G 427

RESULT 13
Q83FF7
ID Q83FF7 PRELIMINARY; PRT; 446 AA.
AC Q83FF7;
DT 01-JUN-2003 (T-EMBLrel. 24, Created)
DT 01-JUN-2003 (T-EMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
DE Enolase (EC 4.2.1.11).
GN Name=eno; OrderedLocuNames=TW7783;
OS Tropheryma whipplei (strain Twist) (Whipple's bacillus).
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Micrococciaceae; Cellulomonadaceae; Tropheryma.
OX NCBI_TaxID=203267;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Twist;
RX MEDLINE=22784088; PubMed=12902375;
RA Recult D., Ogata H., Audic S., Robert C., Suhre K., Drancourt M.,
RA Claverie J.-M.
RT "Tropheryma whipplei Twist: a human pathogenic Actinobacteria with a
RT reduced genome.";
RL Genome Res. 13:1800-1809(2003).
CC -I- CATALYTIC ACTIVITY: 2-phospho-D-glycerate = phosphoenolpyruvate +
CC H(2)O.
CC -I- COPACTOR: Magnesium is required for catalysis and for stabilizing
CC the dimer (By similarity).
CC -I- PATHWAY: Glycolysis.
CC -I- SUBUNIT: Homodimer (By similarity).
CC -I- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -I- SIMILARITY: Belongs to the enolase family.
DR EMBL; AE016852; AA044880.1; --
DR HSSP; Q9NDH8; 10EP.
DR GO; GO:0000015; C:phosphopyruvate hydratase complex; IEA.
DR GO; GO:0016829; F:lyase activity; IEA.
DR GO; GO:0000287; F:magnesium ion binding; IEA.
DR GO; GO:0004634; F:phosphopyruvate hydratase activity; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR000941; Enolase.
DR Pfam; PF00113; Enolase_C; 1.
DR Pfam; PF03952; Enolase_N; 1.
DR PRINTS; PR00148; ENOLASE.
DR PRODOM; PD000902; Enolase; 1.
DR TIGRFAMS; TIGR01060; eno; 1.
DR PROSITE; PS00164; ENOLASE; 1.
KW Complete proteome; Glycolysis; Lyase; Magnesium.
SQ SEQUENCE 446 AA; 48141 MW; 786E73F99F94A6DA CRC64;

Query Match      66.6%; Score 1436; DB 2; Length 446;
Best Local Similarity 67.9%; Pred. No. 4.1e-77;
Matches 286; Conservative 51; Mismatches 82; Indels 2; Gaps 2;

QY 7 VFAREILDSRGNPTVEAEVFLDDGSHGVAGVPSGASTGVHAEHLRD-GGDRLYLGKGVLK 65
Db 25 VHARQILDSRGTPTEVEVTELDGSMGRSAVPSGASTGAFENAEHLRDQDNNEVILGKGVTR 84
QY 66 AVENVNEEIGDELAGLEADDQRLIDEAMIKLDGTANKSRILGANAILGVSMVAKAAADSA 125
Db 85 AVRSVNSEIAPVLIGFDQVGLDHRMIELDGTNNKSRILGANAILGVSLASAAARAA 144
QY 126 GLPLFRYIGGNPAHVLFPVPMNIIINGGAHDSGVVDVQEFMIAPIGAEFTFSEALRNGAEVY 185
Db 145 DLSLFRYLGSPSSRILPVPMMNIIINGGAHDTGVDIOEFMILPVGARFSFSESLRWGCEVY 204
QY 186 HALKSUVKEKGLSTGLDGGFPAPSGVSTREALDLIVKAIEKAGFTPGKDIALLDVASS 245
Db 205 HSLKVQLRESGLSGLDGGFPALRSNRALTDLILSAIEKAGFSPGIDIVLALDIAAS 264
QY 246 EFKE-DGTYHFEFGQHSAAEMANVYAEIVDAYPTVSTEDPLOEDDWEYTNLTATIGDKV 304
Db 265 EFCAPGHYRFEKDTSDLSIYEGLLSSYPLVSTIEDPLDQDDWEGYRTLTTHIGDRV 324
QY 305 QIVGDDDFVTNPNRLKEGIAKKAANSILVKVNOIGTLTETFDVADMAHRAAGYTSMMSHRS 364
Db 325 QIVGDDDLFVNTSRLSRGSIQSGVANSILVKVNOIGTLTETFDVADMAAKGGYTAVALSHRS 384
QY 365 GETEDTTIADLAVALNCQIKTGAPARSDRVAKYNQLLRLEQLLGDAGVYAGRSAPFRFQ 424
Db 385 GETEDTTIADMAVATNCQIKTGAPARGERIAKYNQLLRLEELKGRSARYAGWLSYPRWQ 444
QY 425 G 425
Db 445 G 445

RESULT 14
ENO_AGRT5
ID ENO_AGRT5 STANDARD; PRT; 424 AA.
AC Q8UFH1;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Enolase (EC 4.2.1.11) (2-phosphoglycerate dehydratase) (2-phospho-D-
DE glycerate hydro-lyase).
GN Name=eno; OrderedLocuNames=Atu1426, AGR_C_2631;
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
OX NCBI_TaxID=176299;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608550; PubMed=11743193; DOI=10.1126/science.1066804;
RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. Sr.,
RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
RA Kutayav T., Levy R., Li M.-J., McClelland E., Palmieri A.,
RA Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D.,
RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,

```



```

FT METAL      242      242      Magnesium (By similarity).
FT METAL      285      285      Magnesium (By similarity).
FT METAL      312      312      Magnesium (By similarity).
SQ SEQUENCE   426 AA; 45245 MW; 6F2E0B4816A22458 CRC64;

Query Match      65.5%; Score 1412.5; DB 1; Length 426;
Best Local Similarity 66.4%; Pred. No. 9.5e-76;
Matches 280; Conservative 59; Mismatches 80; Indels 3; Gaps 2;

QY 1 VAEIMHVFAREIILDSRGNPTVEAEVFLDDGSHGVAGVPSGASTGVHEAHELRDGGD-RYL 59
   :||: : ||||| ||||| ||||| :||: ||||| ||||| ||||| ||||| |||||
Db 1 MTEIVDIIAREIILDSRGNPTVEVDVLEDAFGRAAVPSGASTGAHEANEKRDGDKARYL 60
   ||| :||: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 60 KGKVLKAVENVNEEIGDELAGLEADDQRLIDEAMIKLDGTANKSRLGANAILGVSMVAK 119
   ||| :||: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 KGKGVQQAADAVNGEIPDALSGVDAEDQRRVDNLMIELDGTNPKARLGNAILGVSLATAK 120
   ||| :||: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 120 AAADSAGLPLFRYIGGPNNAHVLVPMNMNIINGNAHDSGVVDQVDFMIADIGAEFFSEALR 179
   ||| :||: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 121 AAASAGLPLYKYVGGYNARVLPPTPMNIIINGGAHADNPIDIQEFMILPTGAKDFREGLR 180
   ||| :||: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 180 NGAEVVHALKSVIKEKGLSTGLDGGGFPSPVSGSTREALDLIVKAIKAGFTPGKDIALA 239
   ||| :||: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 181 MGAEIFHALKALKADAGHNTNVGDEGGFAPNLASAEAAALDFIVKAGEKAGTKAGDDFVLG 240
   ||| :||: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 240 LDVASSEFFKDGTYHFEGGQHS--AAEMANVYAEIVDAYPIVSTIEDPLQEDDWEGYTNLT 297
   ||| :||: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 241 LDVASTEFFKNGKYELESGEKSUDPAAAVDYLAGLVAKFPILTIEDCGMAEDDPDGWKLLT 300
   ||| :||: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 298 ATIGDKVQIVGDDFFVTNPERLKEGIAKKAANSILVKVNOIGLITETTFDAVDMAHRAGYT 357
   ||| :||: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 301 DTLGKKVQLVGDDDLFVTNPKRLQMLDKGLANSILVKVNQIGTILSETIDAVELAHRHGYT 360
   ||| :||: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 358 SSMSHRSGETEDTTIADLAVALNCQIKTGAPAPSRVAKYNOLLRIEQLLDAGVYAGR 417
   ||| :||: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 361 SVMSHRSGETEDTTIADLAVALNCQIKTGLSARSDRTAKYNOLLRIQEMLDDOGVYAGR 420
   ||| :||: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 418 SA 419
   :|
Db 421 AA 422

```

Search completed: September 29, 2005, 12:13:47
Job time : 89.5 secs

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 29, 2005, 09:39:18 ; Search time 38.5 Seconds
(without alignments)
824.048 Million cell updates/sec

Title: US-10-728-947-4
Perfect score: 2155
Sequence: 1 VAEIMHVPAREILDSRGNT.....QLLGAGVYGRSAFFRFG 425

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2.6/prodata/1/iaa/5A COMB.pep.*
2: /cgn2.6/prodata/1/iaa/5B COMB.pep.*
3: /cgn2.6/prodata/1/iaa/6A COMB.pep.*
4: /cgn2.6/prodata/1/iaa/6B COMB.pep.*
5: /cgn2.6/prodata/1/iaa/PCUS COMB.pep.*
6: /cgn2.6/prodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2155	100.0	425	US-09-860-768-4	Sequence 4, Appli
2	2151	99.8	425	US-09-860-768-2	Sequence 2, Appli
3	1357.5	63.0	432	US-09-302-540-12060	Sequence 12060, A
4	1306.5	60.6	434	US-09-710-279-2222	Sequence 2222, Ap
5	1306.5	60.6	440	US-09-134-001C-5641	Sequence 5641, Ap
6	1280.5	59.4	434	US-09-830-217-22	Sequence 22, Appl
7	1280.5	59.4	434	US-10-278-946-22	Sequence 22, Appl
8	1278.5	59.3	444	US-09-252-991A-21020	Sequence 21020, A
9	1256.5	58.3	434	US-09-583-110-3817	Sequence 3817, Ap
10	1240	57.5	398	US-09-107-532A-3810	Sequence 3810, Ap
11	1231.5	57.1	444	US-09-328-352-7269	Sequence 7269, Ap
12	1226.5	56.9	420	US-08-847-065-19	Sequence 19, Appl
13	1226.5	56.9	420	US-09-829-382-19	Sequence 19, Appl
14	1222	56.7	459	US-09-489-039A-10061	Sequence 10061, A
15	1201.5	55.8	445	US-09-540-236-2354	Sequence 2354, A
16	1186.5	55.1	456	US-09-543-681A-6090	Sequence 6090, Ap
17	1168.5	54.2	434	US-09-438-185A-802	Sequence 802, App
18	1131.5	52.5	407	US-09-107-433-4986	Sequence 4986, Ap
19	1097	50.9	443	US-09-949-016-8359	Sequence 8359, Ap
20	1095	50.8	434	US-09-949-016-6153	Sequence 6153, Ap
21	942.5	43.7	452	US-09-198-452A-853	Sequence 853, App
22	902	41.9	311	US-08-961-083-184	Sequence 184, App
23	902	41.9	311	US-09-536-784-184	Sequence 184, App
24	651.5	30.2	301	US-09-248-796A-17053	Sequence 17053, A
25	504.5	23.4	177	US-08-446-920-11	Sequence 11, Appl
26	458	21.3	148	US-09-248-796A-17054	Sequence 17054, A
27	265	12.3	113	US-09-513-999C-5420	Sequence 5420, Ap

ALIGNMENTS

RESULT 1

US-09-860-768-4
; Sequence 4, Application US/09860768
; Patent No. 6713289
; GENERAL INFORMATION:
; APPLICANT: Mockel, Bettina
; APPLICANT: Preferlie, Walter
; APPLICANT: Hermann, Thomas
; APPLICANT: Pohler, Alfred
; APPLICANT: Kalinowski, Jorn
; APPLICANT: Bathe, Brigitte
; TITLE OF INVENTION: New Nucleotide Sequences that Code for the Eno Gene
; FILE REFERENCE: 21123/278404
; CURRENT APPLICATION NUMBER: US/09/860,768
; CURRENT FILING DATE: 2001-05-21
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 425
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-860-768-4

Query Match	100.0%;	Score 2155;	DB 4;	Length 425;
Best Local Similarity	100.0%;	Pred. No. 1.8e-190;		
Matches 425;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	VAEIMHVPAREILDSRGNTPTVEAEVFLDDGSHGVAGVPSGASTGVHAEHLRDGDDRYLG	60	
Db	1	VAEIMHVPAREILDSRGNTPTVEAEVFLDDGSHGVAGVPSGASTGVHAEHLRDGDDRYLG	60	
Qy	61	KGVLKAVENVNEETGDELAGEADDQRLIDEAMI KLDGTANKSRLGANAILGVSMVAKA	120	
Db	61	KGVLKAVENVNEETGDELAGEADDQRLIDEAMI KLDGTANKSRLGANAILGVSMVAKA	120	
Qy	121	AADSAGILPLFRYIGGPNNAHLVPPMMNI INCGAHADSGVDVQVEFMIAPIGAETTFSEALRN	180	
Db	121	AADSAGILPLFRYIGGPNNAHLVPPMMNI INCGAHADSGVDVQVEFMIAPIGAETTFSEALRN	180	
Qy	181	GAEVYHALKSVIKKGLSTGLGDEGGFAPSVGSTREALDLILVKAIEKAGFTPGKDIALAL	240	
Db	181	GAEVYHALKSVIKKGLSTGLGDEGGFAPSVGSTREALDLILVKAIEKAGFTPGKDIALAL	240	
Qy	241	DVASSEFFKDGTYHFEGGQHSAAEMANVYAEVLVDAYPIVSTIEDPLQEDDWEQYTNLTATI	300	
Db	241	DVASSEFFKDGTYHFEGGQHSAAEMANVYAEVLVDAYPIVSTIEDPLQEDDWEQYTNLTATI	300	
Qy	301	GDVKQIVGDDDFVTNPERLKEGIAKKAANSILVKVNOIGTLTETTFDAVDMAHRAGYTSM	360	
Db	301	GDVKQIVGDDDFVTNPERLKEGIAKKAANSILVKVNOIGTLTETTFDAVDMAHRAGYTSM	360	

QY 361 SHRSGETEDTTIADLAVALNCGQIKTGAPARSDRVAKYNOLLRIEQLLDGAGVYAGRSF 420
Db |||||
QY 361 SHRSGETEDTTIADLAVALNCGQIKTGAPARSDRVAKYNOLLRIEQLLDGAGVYAGRSF 420
Db |||||
QY 421 PRFQ 425
Db |||||
QY 421 PRFQ 425
Db |||||
RESULT 2
US-09-860-768-2
; Sequence 2, Application US/09860768
; Patent No. 6713289
; GENERAL INFORMATION:
; APPLICANT: Mockel, Bettina
; APPLICANT: Pfefferle, Walter
; APPLICANT: Hermann, Thomas
; APPLICANT: Pohler, Alfred
; APPLICANT: Kalinowski, Jörn
; APPLICANT: Bathe, Brigitte
; TITLE OF INVENTION: New Nucleotide Sequences that Code for the Eno Gene
; FILE REFERENCE: 21123/278404
; CURRENT APPLICATION NUMBER: US/09/860,768
; CURRENT FILING DATE: 2001-05-21
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 425
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-860-768-2

Query Match 99.8%; Score 2151; DB 4; Length 425;
Best Local Similarity 99.8%; Pred. No. 4.3e-190;
Matches 424; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 VAEIMHVFAREILDSRGNPTVEAEVFLDDGSHGVAGVPSGASTGVHEAHELDRGGRYLG 60
Db 1 VAEIMHVFAREILDSRGNPTVEAEVFLDDGSHGVAGVPSGASTGVHEAHELDRGGRYLG 60
QY 61 KGVLKAVENNVNEEIGDELAGLEADDQRLIDEAMIKLDGTANKSRLLGANAILGVSMVAKA 120
Db 61 KGVLKAVENNVNEEIGDELAGLEADDQRLIDEAMIKLDGTANKSRLLGANAILGVSMVAKA 120
QY 121 AADSAGLPLFRYIGGPNNAHVLPPVMNIIINGGAHDSGVVDVQEFMTAPIGAETTFSEALRN 180
Db 121 AADSAGLPLFRYIGGPNNAHVLPPVMNIIINGGAHDSGVVDVQEFMTAPIGAETTFSEALRN 180
QY 181 GAEVYHALKSVIKEKGLSTGLGDEGGFAPSVGSTREALDLIVKAIKAGFTPGKDIALAL 240
Db 181 GAEVYHALKSVIKEKGLSTGLGDEGGFAPSVGSTREALDLIVKAIKAGFTPGKDIALAL 240
QY 241 DVASSEFFKDGTVHFEGGQHSAAEMANVYAEVLVDAYPIVSIEDPLQEDDMEGYTNLTATI 300
Db 241 DVASSEFFKDGTVHFEGGQHSAAEMANVYAEVLVDAYPIVSIEDPLQEDDMEGYTNLTATI 300
QY 301 GDKVQIVGDDFFVTNPERLKEGTAKAANSILVKVNOIGTLTETTFDAVDMHRAAGTYSNM 360
Db 301 GDKVQIVGDDFFVTNPERLKEGTAKAANSILVKVNOIGTLTETTFDAVDMHRAAGTYSNM 360
QY 361 SHRSGETEDTTIADLAVALNCGQIKTGAPARSDRVAKYNOLLRIEQLLDGAGVYAGRSF 420
Db 361 SHRSGETEDTTIADLAVALNCGQIKTGAPARSDRVAKYNOLLRIEQLLDGAGVYAGRSF 420
QY 421 PRFQ 425
Db |||||
QY 421 PRFQ 425
Db |||||

RESULT 3
US-09-902-540-12060
; Sequence 12060, Application US/09902540

; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 12060
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-12060

Query Match 63.0%; Score 1357.5; DB 4; Length 432;
Best Local Similarity 63.8%; Pred. No. 8.4e-117;
Matches 271; Conservative 59; Mismatches 90; Indels 5; Gaps 2;
QY 1 VAEIMHVFAREILDSRGNPTVEAEVFLDDGSHGVAGVPSGASTGVHEAHELDRGG-DRYL 59
Db 1 MTEISQILAREVLDSRGNPTVEAEVQLAGGARAAVPSGASTGEHEAIELRDGDKHRYL 60
QY 60 KGVLKAVENNVNEEIGDELAGLEADDQRLIDEAMIKLDGTANKSRLLGANAILGVSMVAK 119
Db 61 KGVLKAVENNVVDLAPALVGMDDAADOFAVDQRMLELDGTATKGLGANAILAVSMAAAR 120
QY 120 AADSAGLPLFRYIGGPNNAHVLPPVMNIIINGGAHDSGVVDVQEFMTAPIGAETTFSEALR 179
Db 121 AADAAGHGLPLYRYGVGVQARTLPVPLMNTLNGAHADTRVDVQEFMVVPAGASSFAEGLR 180
QY 180 NGAEVYHALKSVIKEKGLSTGLGDEGGFAPSVGSTREALDLIVKAIKAGFTPGKDIALA 239
Db 181 WGAEVYHALKSVIKEKGLSTGLGDEGGFAPSVGSTREALDLIVKAIKAGFTPGKDIALA 240
QY 240 LDVASSEFFKDGTVHF-EGGQHSAAEMANVYAEVLVDAYPIVSIEDPLQEDDMEGYTN 295
Db 241 LDVASSEFFKDGSKKYLKGEKGYDSTGLLEYRYGLSERYPIISIEDGMAEDDMEGWKK 300
QY 296 LTATTGDKVQIVGDDFFVTNPERLKEGTAKAANSILVKVNOIGTLTETTFDAVDMHRAAG 355
Db 301 LTDALGSRMQLVGGDDLFTVNVERLGRGIETGTANSILVKVNOIGTLTETTFDAVDMHRAAG 360
QY 356 YTSMSHRSRGETEDTTIADLAVALNCGQIKTGAPARSDRVAKYNOLLRIEQLLDGAGVYA 415
Db 361 YTSMSHRSRGETEDTTIADLAVALNCGQIKTGAPARSDRVAKYNOLLRIEQLLDGAGVYA 420
QY 416 GRSF 420
Db |||||
QY 421 GRSF 425
Db |||||

RESULT 4
US-09-710-279-2222
; Sequence 2222, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PUS480US
; CURRENT APPLICATION NUMBER: US/09/710,279
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2222
; LENGTH: 434
; TYPE: PRT

; ORGANISM: Artificial Sequence		
; FEATURE:		
; OTHER INFORMATION: Description of Artificial Sequence: synthetic		
; OTHER INFORMATION: amino acid sequence		
US-09-710-279-2222		
QY	Query Match	60.6%; Score 1306.5; DB 4; Length 434;
Db	Best Local Similarity	60.8%; Pred. No. 4.4e-112;
	Matches 259; Conservative	64; Mismatches 94; Indels 9; Gaps 3;
QY	4 IMHVAREILDSRGNTPTVEAEVFLDDGSHGVAGVPSGASTGVHEAHELDRGG--DRYLKGG	62
Db	4 ITDVYAREVLDSRGNTPTVEAEVFLDDGSHGVAGVPSGASTGVHEAHELDRGG--DRYLKGG	63
QY	63 VLKAVENNVNEEIGDELAGLEAD--DQRLIDEAMIKLDGTANKSRILGANAILGVSMVAKA	120
Db	64 VTKAVENNVNEMIAPEIVEGEFSVLDQVSDIKMMIQLDGTHNKGKLGANAILGVSIIVARA	123
QY	121 AADSAGLPLFRYIGGNNAHVLPVPMNIIINGGAHSDGVVDQEFMTAPIGAETTFSEALRN	180
Db	124 AADLLGQPLYKYLGGFNGKQLPVPMMNIVNGGSHSDAPIAFQEFMILPVGAESFKESLRW	183
QY	181 GAENVHALKSVIKEKGLSTGLGDEGGFAPSGVSTREALDLIVKAIKAGFTPGKDIALAL	240
Db	184 GAEIFHNLSILSERGLETAVGDEGGFAPRFEGETDAVETIIKAIKAGYKPGEDVFLGF	243
QY	241 DVASSEFFKDGTYHF-----EGQHSAAEMANVYAEVLVDAYPIVSIEDPLQDDWEGYT	294
Db	244 DCASSEFFYENGVDYTKFEGEHGAKSAARQVDYBELIGKYPIIITIEDGMDENDWEGWK	303
QY	295 NLTTATIGDKVQIVGDDFFVTNPERLKEGIAKKAANSILVKVNOIGTLTETTFDAVDMAHRA	354
Db	304 QLTDRIGDKVQLVGDDLFVTNTEILSKGIEQGIGNSILIKVNOIGTLTETTFDAIEMAQA	363
QY	355 GYTSMMSHRSGETETTTIADLAVALNCGQIKTGAPARSDRVAKYNOLLRIEQLLGDAGVY	414
Db	364 GYTAVVSHRSGETETTTIADIAVATNAGQIKTGSLSRTDRIAKYNOLLRIEDELSETAKF	423
QY	415 AGRSAF 420	
Db	424 EGKSP 429	
RESULT 5		
US-09-134-001C-5641		
; Sequence 5641, Application US/09134001C		
; Patent No. 6380370		
; GENERAL INFORMATION:		
; APPLICANT: Lynn Doucette-Stamm et al		
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS		
; FILE REFERENCE: CTC-007		
; CURRENT APPLICATION NUMBER: US/09/134,001C		
; CURRENT FILING DATE: 1998-08-13		
; PRIOR APPLICATION NUMBER: US 60/064,964		
; PRIOR FILING DATE: 1997-11-08		
; PRIOR APPLICATION NUMBER: US 60/055,779		
; PRIOR FILING DATE: 1997-08-14		
; NUMBER OF SEQ ID NOS: 5674		
; SEQ ID NO 5641		
; LENGTH: 440		
; TYPE: PRT		
; ORGANISM: Staphylococcus epidermidis		
US-09-134-001C-5641		
QY	Query Match	60.6%; Score 1306.5; DB 3; Length 440;
Db	Best Local Similarity	60.8%; Pred. No. 4.5e-112;
	Matches 259; Conservative	64; Mismatches 94; Indels 9; Gaps 3;
QY	4 IMHVAREILDSRGNTPTVEAEVFLDDGSHGVAGVPSGASTGVHEAHELDRGG--DRYLKGG	62
Db	10 ITDVYAREVLDSRGNTPTVEAEVFLDDGSHGVAGVPSGASTGVHEAHELDRGG--DRYLKGG	69

RESULT 6			
US-09-830-217-22			
; Sequence 22, Application US/09830217			
; Patent No. 6521441			
; GENERAL INFORMATION:			
; APPLICANT: Human Genome Sciences, Inc.			
; TITLE OF INVENTION: Staphylococcus aureus Genes and Polypeptides			
; FILE REFERENCE: PB461PCT			
; CURRENT APPLICATION NUMBER: US/09/830,217			
; CURRENT FILING DATE: 2001-04-24			
; PRIOR APPLICATION NUMBER: PCT/US99/06199			
; PRIOR FILING DATE: 1999-03-18			
; PRIOR APPLICATION NUMBER: 60/080,296			
; PRIOR FILING DATE: 1998-04-01			
; PRIOR APPLICATION NUMBER: 60/084,674			
; PRIOR FILING DATE: 1998-05-07			
; NUMBER OF SEQ ID NOS: 22			
; SOFTWARE: PatentIn Ver. 2.0			
; SEQ ID NO 22			
; LENGTH: 434			
; TYPE: PRT			
; ORGANISM: Staphylococcus aureus			
US-09-830-217-22			
Query Match 59.4%; Score 1280.5; DB 4; Length 434;			
Best Local Similarity 59.1%; Pred. No. 1.1e-109;			
Matches 253; Conservative 68; Mismatches 94; Indels 13; Gaps 4;			
Qy	4	IMHVAREILDSRGNTPTVEAEVFLDDGSHGVAGVPSGASTGVHEAHELDRGG--DRYLKGG	62
Db	4	ITDVYAREVLDSRGNTPTVEAEVFLDDGSHGVAGVPSGASTGVHEAHELDRGG--DRYLKGG	63
Qy	63	VLKAVENNVNEEIGDELAGLEAD---DQRLIDEAMIKLDGTANKSRILGANAILGVSMAVA	118
Db	64	VTKAVENNVNEITAPEI--IEGEFSVLDDQVSDIKMMIALDGTENKGLGANAILGVSIAVA	121
Qy	119	KAAADSAGLPLFRYIGGNNAHVLPVPMNIIINGGAHSDGVVDQEFMTAPIGAETTFSEAL	178
Db	122	RAAADLLGQPLYKYLGGFNGKQLPVPMMNIVNGGSHSDAPIAFQEFMILPVGATTTFKESL	181
Qy	179	RGAENVHALKSVIKEKGLSTGLGDEGGFAPSGVSTREALDLIVKAIKAGFTPGKDIAL	238
Db	182	RWGTEIFHNLSILSQRLGLETAVGDEGGFAPKPEGTEDAVETIIQAIKAGYKPGEEVFL	241


```
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; FILE REFERENCE: PATH00-07A
; CURRENT APPLICATION NUMBER: US/09/583,110
; CURRENT FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/107,433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 3817
; LENGTH: 434
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
; US-09-583-110-3817

Query Match      58.3%; Score 1256.5; DB 4; Length 434;
Best Local Similarity 58.5%; Pred. No. 1.8e-107;
Matches 250; Conservative 65; Mismatches 101; Indels 11; Gaps 4;

QY 4 IMVFAREILDSRGNPTVEAEVFLDDGSHGVAGVPSGASTGVVHAHELRLDGG-DRYLKKG 62
DB 4 ITDVYAREVLDSRGNPTLEVEVYTESGAFGRGMVPSGASTGEHEAVELRDGDKSRYYGLG 63
QY 63 VLKAVENNEEIGDELGLADDOORLIDEAMIKLDGTANKSRILGANAILGVSNAVAKAAA 122
DB 64 TQKAVDNVNNIIAEIIGYDVRQOQALDRAMIALDGTNPNGKILGANAILGVSIAVARAAA 123
QY 123 DSAGLPLFRYIGGNPAHLVPPVPMNIIINGGAHDSGVVDVDFMIAPITGAFTSEALRNGA 182
DB 124 DYLEIPLYSLGGFNKVLPTPMNIIINGGSHSDAPLAFQEFMLLPVGAFTFEALRYGA 183
QY 183 EVTHALKSVIKKGLSTGLDGGEPAPSVGSTREALDLIVKAIEKAGFTPGKDIALADV 242
DB 184 EIFHALKILKSLGLETAVDEGGFAPRPEGTEDGVETILAAEAGYVPGKDVFGFDC 243
QY 243 ASSEFEK-----DGYTHFEG---GQHSAAEMANVYAEILDVAPIVSIEDPQEDDMEGY 293
DB 244 ASSEFYDKERKVYDYT-KFEFGAAVRTSAEQIDYLEELVKNKYPITIEDGMDENDWDGW 302
QY 294 TNLATIGDKVQIVGDDFFVYTNPERLKEGIAKKAANSILVKVNIQIGTLTETFDVADMAHR 353
DB 303 KALTERLGKVKVLVGDDFFVYTNLDYLRGLQEGRAANSILIKVNIQIGTLTETFEAIEWAKE 362
QY 354 AGTSMMSHRSGETEDTTIADLAVALNCGOIKTGAPARSDRVAKYNQILLRIEQLLDGAGV 413
DB 363 AGYTAVVSHRSGETEDSTIADIAVATNAGQIKTGSLSRDRIAKYNQILLRIEQLGVAE 422
QY 414 YACRSAP 420
DB 423 YRLKXSF 429

RESULT 10
US-09-107-532A-3810
; Sequence 3810, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO...
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESS: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
```

```
; MEDIUM TYPE: CD/ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 3810:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 398 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (B) LOCATION 1...398
; SEQUENCE DESCRIPTION: SEQ ID NO: 3810:
US-09-107-532A-3810

Query Match      57.5%; Score 1240; DB 4; Length 398;
Best Local Similarity 62.1%; Pred. No. 5.3e-106;
Matches 242; Conservative 59; Mismatches 83; Indels 6; Gaps 3;

QY 37 VPSGASTGVHEAHELRLDGGD-RYLKGVLKAVENNEEIGDELGLADDOORLIDEAMIK 95
DB 3 VPSGASTGEVEAEVLDRDGGDKARYGGKGVTKAVDNVNNIIAEIIGYDVRQOQALDKAMIA 62
QY 96 LDGTANKSRILGANAILGVSNAVAKAAADSAGLPLFRYIGGNPAHLVPPVPMNIIINGGAH 155
DB 63 LDGTNPNGKILGANAILGVSIAVARAAADYLEVPLHYLGGFNKVLPTPMNIIINGGSHA 122
QY 156 DSGVDVQEFMIAPIGAEFTSEALRNGAEVYHALKSVIKKGLSTGLDGGEPAPSVGSTR 215
DB 123 DNSIDPQEFMIMPVGAPTFKEALRMGAEVFHALASILKARGLATSVGDEGFPANLGSNE 182
QY 216 EALDLIVKAIEKAGFTPGKDIALADVASSERF--KDGTYHFEF--GQHSAAEMANVYA 270
DB 183 EGFEVIEAEKAGYVPGKDVVLADAAASEFVDKEKGVYVLADSGEGEKTDEMIKFEY 242
QY 271 ELVDAYPIVSIEDPQEDDMEGYTNLTATIGDKVQIVGDDFFVYTNPERLKEGIAKKAANS 330
DB 243 ELVSKYPIISIEDGLDENDWDGPKLTDVLGDKVQLVGDDLFVYTNQKLSGIEKGIANS 302
QY 331 ILVKVNIQIGTLTETFDVADMAHRAGYTSMMSHRSGETEDTTIADLAVALNCGOIKTGAPA 390
DB 303 ILIKVNIQIGTLTETFEAIEWAKEAGYTVAVVSHRSGETEDSTIADIAVATNAGQIKTGSL 362
QY 391 RSDRVAKYNQILLRIEQLLDGAGVYAGRSAP 420
DB 363 RTDRIAKYNQILLRIEQLGVAEYKGLKSF 392

RESULT 11
US-09-328-352-7269
; Sequence 7269, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
```

```

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: GTC99-03PA
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 7269
; LENGTH: 444
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-7269

```

```

Query Match      57.1%; Score 1231.5; DB 4; Length 444;
Best Local Similarity 59.3%; Pred. No. 3.8e-105;
Matches 253; Conservative 65; Mismatches 100; Indels 9; Gaps 5;

QY 1 VAEIMHVAFRILDSRGNPTVEAEVFLDDGSHGVAGVPSGASTGVHAEHLRDGG-DRYL 59
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 16 MSQIVDIRAREILDSRGNPTIEADVILESGVGRACAPSGASTGSRLEALRDGDKSYL 75
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 60 GKGVLKAVENNEEIGDELGLAGLEADDQRLIDEAMIKLDGTANKSRILGANAILGVSMVAK 119
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 76 GKGVRTAVQNVNNSIHLLVQSVFEQKALDERKIAFDGTENKSKLGNATLAVSLAAAH 135
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 120 AAADSAGLPLFRYIG---GPNHVLVPMNININGGAHADSGVDVQSFMIAPIGAEFTSE 176
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 136 AAAAEQKLPFLQYIANLRGQTLTMPVPMNINLNGGAHADNTVDIQEFMIEPVGFTSFAE 195
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 177 ALRNGAEVYHALKSIVIEKGLSTGLDGGFAPSVGSTRALDLIVKAEKAGFTPGKOI 236
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 196 ALRGAEEVHSLKSVLKQGLNTAVDGGFAPNLRSEEAITVILQAEQTYKAGSDI 255
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 237 ALALDVASSBFFKDGTYHFEG-GQHS--AAEMANVYAEVDVPIVSIEDPLOEDDWEY 293
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 256 MLALDCASSBFFYKNGQVILSEGNKSFSTNQPADYLAGLVKQPIIISIEDGLDESDEWEG 315
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 294 TNLTAITGDKVQVIGDDFFVTNPERLKEGIAKKAANSILVKVNIQIGTLTTFDAVDMHR 353
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 316 SYLTSILGDKIQLVGDGLFVTNPKILQRIQDEKGVNSILIKYNIQIGTLTTLDAIYLAKA 375
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 354 AGTSMMSHRSGETEDTTIADLAVALNCGOIKTGAPARSDRVAKYNQLLRIEQLLDGAV 413
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 376 NGYTVIHSRSGEETDSTIADLAVGTAAGQIKTGSICRSRVSKYNQLLRIEEL--TKAV 433
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 414 YAGRSF 420
Db : : : : :
QY 434 YRGKAEF 440
Db : : : : :

```

```

RESULT 12
US-08-847-065-19
; Sequence 19, Application US/08847065
; Patent No. 6245335
; GENERAL INFORMATION:
; APPLICANT: Masure, H. Robert
; APPLICANT: Masure, H. Robert
; APPLICANT: Rosenow, Carsten I.
; APPLICANT: Tuomanen, Elaine
; APPLICANT: Wizemann, Theresa M.
; TITLE OF INVENTION: CHOLINE BINDING PROTEINS FOR
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David A. Jackson, Esq.
; STREET: 411 Hackensack Ave, Continental Plaza, 4th
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

```

```

; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/847,065
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-158
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 420 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: B. subtilis
US-08-847-065-19

```

```

Query Match      56.9%; Score 1226.5; DB 3; Length 420;
Best Local Similarity 58.7%; Pred. No. 1e-104;
Matches 244; Conservative 61; Mismatches 100; Indels 11; Gaps 4;

QY 15 SRGNPTVEAEVFLDDGSHGVAGVPSGASTGVHAEHLRDGG-DRYLKGVLKAVENNEE 73
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1 SRGNPTLEVEVYTESAGFQGMVPSGASTGEHAEVLRDGDKSRYGGLGTQKAVDNVNI 60
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 74 IGDELAGLEADDQRLIDEAMIKLDGTANKSRILGANAILGVSMVAKAAADASAGLPLFRYI 133
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 61 IADSIIGVDRVQQAIDRAMIALDGTGPNKLGANAILGVSIARAFADADYLEIPLYSYL 120
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 134 GGNHVLVPMNININGGAHADSGVDVQEFMIAPIGAEFTSEALRNGAEVYHALKSVIK 193
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 121 GGFNTKVLPTPMNININGGSHSDAPAFQEFMILPVGAPTFKTLRYGAEIHALKKILK 180
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 194 EKSLSTGLDGGFAPSVGSTRALDLIVKAEKAGFTPGKDIALALDVASSEFFK--- 249
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 181 SRGLETAVDGGFAPRFEGETEDGVETILAAIEAGYVPGKDVFLGFCASSEFYDKERK 240
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 250 ---DGYHFEG---GQHSAAEMANVYAEVDVPIVSIEDPLOEDDWEYTNLTATIGDKV 304
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 241 VYDYT-KFEGEAAVRTSAEQIDYLEELVNKYPITIEDGMDENDWDGKALTERLGKV 299
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 305 QIVGDDFFVTNPERLKEGIAKKAANSILVKVNIQIGTLTTFDAVDMHRAGYTSMMSHRS 364
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 300 QLVGDDFFVTNIDYLRGKGAANSILIKVNIQIGTLTTFEAIEMAKEAGYTAVVSHRS 359
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 365 GETEDTTIADLAVALNCGOIKTGAPARSDRVAKYNQLLRIEQLLDGAVYAGRSF 420
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 360 GETEDSTIADIATNAGQIKTGSLSRTDRIAKYNQLLRIEQLLDGAVYAGRSF 415
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

RESULT 13
US-09-829-382-19
; Sequence 19, Application US/09829382
; Patent No. 6784164
; GENERAL INFORMATION:
; APPLICANT: Masure, H. Robert
; APPLICANT: Rosenow, Carsten I.
; APPLICANT: Tuomanen, Elaine
; APPLICANT: Wizemann, Theresa M.
; TITLE OF INVENTION: CHOLINE BINDING PROTEINS FOR
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David A. Jackson, Esq.
; STREET: 411 Hackensack Ave, Continental Plaza, 4th

```


Query Match		55.8%;	Score 1201.5;	DB 4;	Length 445;
Best Local Similarity		57.4%;	Pred. No. 2.3e-102;		
Matches 248;		Conservative 64;	Mismatches 111;	Indels 9;	Gaps 4;
Qy	1	VAEIMHVFAREILDSRGNPTVEAEVFLDDGSHGVAGVPSGASTGVHEAHELDDGG-DRYL	59		
Db	16	IVEIKDIIAREILDSRGNPTIEADVILANGIRGPAAPSGASTGSRALERDDGTSRYL	75		
Qy	60	GKVLKAVENNEBIGELAGLEADDORLIDEAMIKLDGTANKSRRLGANAILGVSMVAK	119		
Db	76	GKVKAVANANSQIRTVLLDKDVTQQTQIDNIIELDGTENKGMGANAMLAVALAK	135		
Qy	120	AAADSAGLPLFRYIG---GPNHVLVPMNIIINGGAHDSGVVDVQEFMIAPIGAETPSE	176		
Db	136	AAALAQSLPLHQYIANLRGQTSLTMPVPMNIIINGGAHADNTVDIOEFMIEPVGFASFSE	195		
Qy	177	ALRNGAEVYHALKSVIKEKGLSTGLDGGFAPSVGSTREALDLIVKAIKAGFTPGKDI	236		
Db	196	ALRAGTEIFHALKSVLKAQGLNTAVGDEGGFAPNLSNEEAITVIMQAIKAGYKAGENI	255		
Qy	237	ALALDVASSEFFKDGTYHFEQGQHSAAE---MANVYAEIVDAYPIVISEDPLEDDWEGY	293		
Db	256	FLALDCAASEFYKNGQYILAGEGNKAFDSQFSYLTSLCNQYPIIISIEDGLDESDEGW	315		
Qy	294	TNLTATIGDKVQIVGDDFFVTNPERLKEGIAKKAANSILVKVNCIGTLTETFDVDMHR	353		
Db	316	AYLTQQLGKKVQLVGDDLFVTNPKILOEGIDKDIANAILKFNQIGTLSETLDAILYAKQ	375		
Qy	354	AGYTSMMSHRSGETEDTTIADLAVALNCGQIKTGAPARSDRVAKYNQLRIEQLLDAGV	413		
Db	376	NGYATVISHRSGETEDSTIADLAVGTAAGQIKTGSLSRDRVAKYNQLRIEQQVRAA--	433		
Qy	414	YAGRSAPPRFQG	425		
Db	434	YRGREFIGURG	445		

Search completed: September 29, 2005, 12:30:44
Job time : 39.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 29, 2005, 12:13:58 ; Search time 1026 Seconds
(without alignments)
171.954 Million cell updates/sec

Title: US-10-728-947-4

Perfect score: 2155

Sequence: 1 VAEIMHVFAREILDSRGNT.....QLLGDAVYAGRSAPFRFG 425

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1846076 seqs, 415116000 residues

Total number of hits satisfying chosen parameters: 1846076

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/2/pubpaa/PCT PUBCOMB.pcp.*
2: /cgn2_6/ptodata/2/pubpaa/PCT NEW PUB.pcp.*
3: /cgn2_6/ptodata/2/pubpaa/US05 NEW PUB.pcp.*
4: /cgn2_6/ptodata/2/pubpaa/US06 PUBCOMB.pcp.*
5: /cgn2_6/ptodata/2/pubpaa/US07 NEW PUB.pcp.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS PUBCOMB.pcp.*
7: /cgn2_6/ptodata/2/pubpaa/US08 NEW PUB.pcp.*
8: /cgn2_6/ptodata/2/pubpaa/US09 PUBCOMB.pcp.*
9: /cgn2_6/ptodata/2/pubpaa/US09A PUBCOMB.pcp.*
10: /cgn2_6/ptodata/2/pubpaa/US09B PUBCOMB.pcp.*
11: /cgn2_6/ptodata/2/pubpaa/US09C PUBCOMB.pcp.*
12: /cgn2_6/ptodata/2/pubpaa/US09 NEW PUB.pcp.*
13: /cgn2_6/ptodata/2/pubpaa/US10A PUBCOMB.pcp.*
14: /cgn2_6/ptodata/2/pubpaa/US10B PUBCOMB.pcp.*
15: /cgn2_6/ptodata/2/pubpaa/US10C PUBCOMB.pcp.*
16: /cgn2_6/ptodata/2/pubpaa/US10D PUBCOMB.pcp.*
17: /cgn2_6/ptodata/2/pubpaa/US10E PUBCOMB.pcp.*
18: /cgn2_6/ptodata/2/pubpaa/US10 NEW PUB.pcp.*
19: /cgn2_6/ptodata/2/pubpaa/US11A PUBCOMB.pcp.*
20: /cgn2_6/ptodata/2/pubpaa/US11 NEW PUB.pcp.*
21: /cgn2_6/ptodata/2/pubpaa/US60 NEW PUB.pcp.*
22: /cgn2_6/ptodata/2/pubpaa/US60 PUBCOMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2155	100.0	425	9 US-09-860-768-4	Sequence 4, Appli
2	2155	100.0	425	16 US-10-728-947-4	Sequence 4, Appli
3	2151	99.8	425	9 US-09-860-768-2	Sequence 2, Appli
4	2151	99.8	425	16 US-10-728-947-2	Sequence 2, Appli
5	2148	99.7	425	9 US-09-738-626-4585	Sequence 4585, Ap
6	2145	99.5	425	16 US-10-781-014-72	Sequence 72, Appli
7	2145	99.5	425	17 US-10-494-836-8	Sequence 8, Appli
8	1913	88.8	425	15 US-10-282-122A-53908	Sequence 53908, A
9	1562	72.5	423	15 US-10-369-493-8340	Sequence 8340, Ap
10	1560.5	72.4	429	15 US-10-282-122A-62133	Sequence 62133, A
11	1553.5	72.1	429	15 US-10-282-122A-64490	Sequence 64490, A

12	1553.5	72.1	439	15	US-10-282-122A-62696	Sequence 62696, A
13	1551.5	72.0	428	14	US-10-156-761-11070	Sequence 11070, A
14	1502.5	69.7	447	15	US-10-282-122A-63682	Sequence 63682, A
15	1422	66.0	423	15	US-10-369-493-11696	Sequence 11696, A
16	1412.5	65.5	426	15	US-10-369-493-16911	Sequence 16911, A
17	1402.5	65.1	423	15	US-10-369-493-10557	Sequence 10557, A
18	1402.5	65.1	426	15	US-10-369-493-17886	Sequence 17886, A
19	1402	65.1	429	15	US-10-369-493-17433	Sequence 17433, A
20	1396	64.8	424	15	US-10-369-493-11819	Sequence 11819, A
21	1394	64.7	420	15	US-10-369-493-14739	Sequence 14739, A
22	1394	64.7	423	15	US-10-369-493-15222	Sequence 15222, A
23	1392.5	64.6	431	15	US-10-282-122A-52760	Sequence 52760, A
24	1377.5	63.9	430	15	US-10-369-493-23299	Sequence 23299, A
25	1373	63.7	431	15	US-10-282-122A-46167	Sequence 46167, A
26	1366.5	63.4	448	15	US-10-282-122A-52925	Sequence 52925, A
27	1362.5	63.2	426	15	US-10-369-493-21730	Sequence 21730, A
28	1362.5	63.2	427	15	US-10-369-493-21006	Sequence 21006, A
29	1359.5	63.1	430	15	US-10-282-122A-60954	Sequence 60954, A
30	1357.5	63.0	430	15	US-10-369-493-19491	Sequence 19491, A
31	1354.5	62.9	431	15	US-10-282-122A-51920	Sequence 51920, A
32	1351.5	62.7	427	15	US-10-282-122A-47804	Sequence 47804, A
33	1347.5	62.5	417	15	US-10-369-493-9097	Sequence 9097, Ap
34	1347.5	62.5	424	15	US-10-369-493-7762	Sequence 7762, Ap
35	1347.5	62.5	426	15	US-10-369-493-4688	Sequence 4688, Ap
36	1347.5	62.5	427	15	US-10-369-493-7447	Sequence 7447, Ap
37	1347.5	62.5	427	15	US-10-282-122A-49282	Sequence 49282, A
38	1342.5	62.3	428	16	US-10-375-266-38	Sequence 38, Appl
39	1342	62.3	429	9	US-09-815-242-5012	Sequence 5012, Ap
40	1342	62.3	432	9	US-09-815-242-10854	Sequence 10854, A
41	1342	62.3	432	15	US-10-282-122A-42554	Sequence 42554, A
42	1334.5	61.9	426	15	US-10-369-493-32	Sequence 32, Appl
43	1332.5	61.8	428	15	US-10-369-493-9970	Sequence 9970, Ap
44	1327.5	61.6	424	15	US-10-369-493-20662	Sequence 20662, A
45	1322	61.3	422	15	US-10-369-493-10435	Sequence 10435, A

ALIGNMENTS

RESULT 1
US-09-860-768-4
; Sequence 4, Application US/09860768
; Patent No. US20020082403A1
; GENERAL INFORMATION:
; APPLICANT: Mockel, Bettina
; APPLICANT: Pfefferle, Walter
; APPLICANT: Hermann, Thomas
; APPLICANT: Pohler, Alfred
; APPLICANT: Kalinowski, Jorn
; APPLICANT: Bathe, Brigitte
; TITLE OF INVENTION: New Nucleotide Sequences that Code for the Eno Gene
; FILE REFERENCE: 21123/278404
; CURRENT APPLICATION NUMBER: US/09/860,768
; CURRENT FILING DATE: 2001-05-21
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 4
; LENGTH: 425
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-860-768-4

Query Match 100.0%; Score 2155; DB 9; Length 425;
Best Local Similarity 100.0%; Pred. No. 6.7e-171;
Matches 425; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	VAEIMHVFAREILDSRGNTVFAEVLDDGSHGVAGVPSGASTGVHAEHLRDGDRYLG	60
DB	1	VAEIMHVFAREILDSRGNTVFAEVLDDGSHGVAGVPSGASTGVHAEHLRDGDRYLG	60
QY	61	KGVLKAVENVNBEIGDELGLADDDORLIDEAMIKLDGTANKSRLLGANAILGVSMVAKA	120
DB	61	KGVLKAVENVNBEIGDELGLADDDORLIDEAMIKLDGTANKSRLLGANAILGVSMVAKA	120

QY 121 AADSAGLPLFRYIGGPNHVLVPMNIIINGGAHADSGVDVQEFMIAPIGAETPSEALRN 180
 DB 121 AADSAGLPLFRYIGGPNHVLVPMNIIINGGAHADSGVDVQEFMIAPIGAETPSEALRN 180
 QY 181 GAEVYHALKSVIKKGLSTGLDGGGFPAPSVGSTREALDLIVKAIKAGFTPGKDIALAL 240
 DB 181 GAEVYHALKSVIKKGLSTGLDGGGFPAPSVGSTREALDLIVKAIKAGFTPGKDIALAL 240
 QY 241 DVASSEFFKQGYHFEFGQHSAAEMANVYAEVLVDAYPIVSIEDPLQEDDWEGYTNLTATI 300
 DB 241 DVASSEFFKQGYHFEFGQHSAAEMANVYAEVLVDAYPIVSIEDPLQEDDWEGYTNLTATI 300
 QY 301 GDKVQIVGDDFFVTNPERLKEGIAKKAANSILVKVNOIGTLTETFDVADVMHAHRAGYTSMM 360
 DB 301 GDKVQIVGDDFFVTNPERLKEGIAKKAANSILVKVNOIGTLTETFDVADVMHAHRAGYTSMM 360
 QY 361 SHRSGETEDTTIADLAVALNCQIKTGAPARSDRVAKYNQLLRIEQLLDGAGVYAGRSAP 420
 DB 361 SHRSGETEDTTIADLAVALNCQIKTGAPARSDRVAKYNQLLRIEQLLDGAGVYAGRSAP 420
 QY 421 PRFQ 425
 DB 421 PRFQ 425

RESULT 2

US-10-728-947-4
 ; Sequence 4, Application US/10728947
 ; Publication No. US2004020394A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Mockel, Bettina
 ; APPLICANT: Pfefferle, Walter
 ; APPLICANT: Hermann, Thomas
 ; APPLICANT: Pohler, Alfred
 ; APPLICANT: Kalinowski, Jorn
 ; APPLICANT: Bathe, Brigitte
 ; TITLE OF INVENTION: New Nucleotide Sequences that Code for the Eno Gene
 ; FILE REFERENCE: 21123/278404
 ; CURRENT APPLICATION NUMBER: US/10/728,947
 ; PRIOR FILING DATE: 2003-12-08
 ; PRIOR APPLICATION NUMBER: US/09/860,768
 ; PRIOR FILING DATE: 2001-05-21
 ; NUMBER OF SEQ ID NOS: 6
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 4
 ; LENGTH: 425
 ; TYPE: PRT
 ; ORGANISM: Corynebacterium glutamicum
 US-10-728-947-4

Query Match 100.0%; Score 2155; DB 16; Length 425;
 Best Local Similarity 100.0%; Pred. No. 6.7e-171;
 Matches 425; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VAEIMHVFAREILDSRGNTVFAEVFLDDGSHGVAGVPSGASTGVHEAHELKGGDRYLG 60
 DB 1 VAEIMHVFAREILDSRGNTVFAEVFLDDGSHGVAGVPSGASTGVHEAHELKGGDRYLG 60
 QY 61 KGVLKAVENNVNEEIGDELAGLEADDQRLIDEAMI KLDGTANKSRLGANAILGVSMVAVAKA 120
 DB 61 KGVLKAVENNVNEEIGDELAGLEADDQRLIDEAMI KLDGTANKSRLGANAILGVSMVAVAKA 120
 QY 121 AADSAGLPLFRYIGGPNHVLVPMNIIINGGAHADSGVDVQEFMIAPIGAETPSEALRN 180
 DB 121 AADSAGLPLFRYIGGPNHVLVPMNIIINGGAHADSGVDVQEFMIAPIGAETPSEALRN 180
 QY 181 GAEVYHALKSVIKKGLSTGLDGGGFPAPSVGSTREALDLIVKAIKAGFTPGKDIALAL 240
 DB 181 GAEVYHALKSVIKKGLSTGLDGGGFPAPSVGSTREALDLIVKAIKAGFTPGKDIALAL 240
 QY 241 DVASSEFFKQGYHFEFGQHSAAEMANVYAEVLVDAYPIVSIEDPLQEDDWEGYTNLTATI 300
 DB 241 DVASSEFFKQGYHFEFGQHSAAEMANVYAEVLVDAYPIVSIEDPLQEDDWEGYTNLTATI 300

DB 241 DVASSEFFKQGYHFEFGQHSAAEMANVYAEVLVDAYPIVSIEDPLQEDDWEGYTNLTATI 300
 QY 301 GDKVQIVGDDFFVTNPERLKEGIAKKAANSILVKVNOIGTLTETFDVADVMHAHRAGYTSMM 360
 DB 301 GDKVQIVGDDFFVTNPERLKEGIAKKAANSILVKVNOIGTLTETFDVADVMHAHRAGYTSMM 360
 QY 361 SHRSGETEDTTIADLAVALNCQIKTGAPARSDRVAKYNQLLRIEQLLDGAGVYAGRSAP 420
 DB 361 SHRSGETEDTTIADLAVALNCQIKTGAPARSDRVAKYNQLLRIEQLLDGAGVYAGRSAP 420
 QY 421 PRFQ 425
 DB 421 PRFQ 425

RESULT 3
 US-09-860-768-2
 ; Sequence 2, Application US/09860768
 ; Patent No. US20020082403A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Mockel, Bettina
 ; APPLICANT: Pfefferle, Walter
 ; APPLICANT: Hermann, Thomas
 ; APPLICANT: Pohler, Alfred
 ; APPLICANT: Kalinowski, Jorn
 ; APPLICANT: Bathe, Brigitte
 ; TITLE OF INVENTION: New Nucleotide Sequences that Code for the Eno Gene
 ; FILE REFERENCE: 21123/278404
 ; CURRENT APPLICATION NUMBER: US/09/860,768
 ; CURRENT FILING DATE: 2001-05-21
 ; NUMBER OF SEQ ID NOS: 6
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 2
 ; LENGTH: 425
 ; TYPE: PRT
 ; ORGANISM: Corynebacterium glutamicum
 US-09-860-768-2

Query Match 99.8%; Score 2151; DB 9; Length 425;
 Best Local Similarity 99.8%; Pred. No. 1.5e-170;
 Matches 424; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VAEIMHVFAREILDSRGNTVFAEVFLDDGSHGVAGVPSGASTGVHEAHELKGGDRYLG 60
 DB 1 VAEIMHVFAREILDSRGNTVFAEVFLDDGSHGVAGVPSGASTGVHEAHELKGGDRYLG 60
 QY 61 KGVLKAVENNVNEEIGDELAGLEADDQRLIDEAMI KLDGTANKSRLGANAILGVSMVAVAKA 120
 DB 61 KGVLKAVENNVNEEIGDELAGLEADDQRLIDEAMI KLDGTANKSRLGANAILGVSMVAVAKA 120
 QY 121 AADSAGLPLFRYIGGPNHVLVPMNIIINGGAHADSGVDVQEFMIAPIGAETPSEALRN 180
 DB 121 AADSAGLPLFRYIGGPNHVLVPMNIIINGGAHADSGVDVQEFMIAPIGAETPSEALRN 180
 QY 181 GAEVYHALKSVIKKGLSTGLDGGGFPAPSVGSTREALDLIVKAIKAGFTPGKDIALAL 240
 DB 181 GAEVYHALKSVIKKGLSTGLDGGGFPAPSVGSTREALDLIVKAIKAGFTPGKDIALAL 240
 QY 241 DVASSEFFKQGYHFEFGQHSAAEMANVYAEVLVDAYPIVSIEDPLQEDDWEGYTNLTATI 300
 DB 241 DVASSEFFKQGYHFEFGQHSAAEMANVYAEVLVDAYPIVSIEDPLQEDDWEGYTNLTATI 300
 QY 301 GDKVQIVGDDFFVTNPERLKEGIAKKAANSILVKVNOIGTLTETFDVADVMHAHRAGYTSMM 360
 DB 301 GDKVQIVGDDFFVTNPERLKEGIAKKAANSILVKVNOIGTLTETFDVADVMHAHRAGYTSMM 360
 QY 361 SHRSGETEDTTIADLAVALNCQIKTGAPARSDRVAKYNQLLRIEQLLDGAGVYAGRSAP 420
 DB 361 SHRSGETEDTTIADLAVALNCQIKTGAPARSDRVAKYNQLLRIEQLLDGAGVYAGRSAP 420
 QY 421 PRFQ 425
 DB 421 PRFQ 425

RESULT 4
US-10-728-947-2
; Sequence 2, Application US/10728947
; Publication No. US20040220394A1
; GENERAL INFORMATION:
; APPLICANT: Mockel, Bettina
; APPLICANT: Pfefferle, Walter
; APPLICANT: Hermann, Thomas
; APPLICANT: Pohler, Alfred
; APPLICANT: Kalinowski, Jörn
; APPLICANT: Bache, Brigitte
; TITLE OF INVENTION: New Nucleotide Sequences that Code for the Eno Gene
; FILE REFERENCE: 21123/278404
; CURRENT APPLICATION NUMBER: US/10728,947
; PRIOR FILING DATE: 2003-12-08
; PRIOR APPLICATION NUMBER: US/09/860,768
; PRIOR FILING DATE: 2001-05-21
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 425
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-10-728-947-2

Query Match 99.8%; Score 2151; DB 16; Length 425;
Best Local Similarity 99.8%; Pred. No. 1.5e-170; Length 425;
Matches 424; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 1 VAEIMHVFAREIILDSRGNTPTVEAEVFLDDGSHGVAGVPSGASTGVHEAHLRGGDRYLG 60
Db 1 VAEIMHVFAREIILDSRGNTPTVEAEVFLDDGSHGVAGVPSGASTGVHEAHLRGGDRYLG 60
Qy 61 KGVLKAVENNEEIGDELADDDQRLIDEAMIKLDGTANKSRGLGANAILGVSMVAKA 120
Db 61 KGVLKAVENNEEIGDELADDDQRLIDEAMIKLDGTANKSRGLGANAILGVSMVAKA 120
Qy 121 AADSAGLPLFRYIGGPNHVLVPMNIIINGGAHDSGVVDVQBFMIAPIGAETTFSEALRN 180
Db 121 AADSAGLPLFRYIGGPNHVLVPMNIIINGGAHDSGVVDVQBFMIAPIGAETTFSEALRN 180
Qy 181 GAEVYHALKSVIKEGLSTGLGDEGGFAPSVGSTREALDLIVKAEKAGTTPGKDIALAL 240
Db 181 GAEVYHALKSVIKEGLSTGLGDEGGFAPSVGSTREALDLIVKAEKAGTTPGKDIALAL 240
Qy 241 DVASSEPFKDGTYHFEGGQSAEMANVYAEVLDAYPIVSIEDPLQEDDWEYTNLTATI 300
Db 241 DVASSEPFKDGTYHFEGGQSAEMANVYAEVLDAYPIVSIEDPLQEDDWEYTNLTATI 300
Qy 301 GDKVQIVGDDFFVTNPERLKEGIAKKAANSILVKVNOIGTITETFDVADMAHRAGYTSM 360
Db 301 GDKVQIVGDDFFVTNPERLKEGIAKKAANSILVKVNOIGTITETFDVADMAHRAGYTSM 360
Qy 361 SHRSGETEDTTIADLAVALNCGQIKTGAPARSDRVAKYNOLLRIEQLLDGAGVYAGRSAP 420
Db 361 SHRSGETEDTTIADLAVALNCGQIKTGAPARSDRVAKYNOLLRIEQLLDGAGVYAGRSAP 420
Qy 421 PRFQG 425
Db 421 PRFQG 425

RESULT 5
US-09-738-626-4585
; Sequence 4585, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO

; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 4585
; LENGTH: 425
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-4585

Query Match 99.7%; Score 2148; DB 9; Length 425;
Best Local Similarity 99.5%; Pred. No. 2.6e-170; Length 425;
Matches 423; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
Qy 1 VAEIMHVFAREIILDSRGNTPTVEAEVFLDDGSHGVAGVPSGASTGVHEAHLRGGDRYLG 60
Db 1 MAEIMHVFAREIILDSRGNTPTVEAEVFLDDGSHGVAGVPSGASTGVHEAHLRGGDRYLG 60
Qy 61 KGVLKAVENNEEIGDELADDDQRLIDEAMIKLDGTANKSRGLGANAILGVSMVAKA 120
Db 61 KGVLKAVENNEEIGDELADDDQRLIDEAMIKLDGTANKSRGLGANAILGVSMVAKA 120
Qy 121 AADSAGLPLFRYIGGPNHVLVPMNIIINGGAHDSGVVDVQBFMIAPIGAETTFSEALRN 180
Db 121 AADSAGLPLFRYIGGPNHVLVPMNIIINGGAHDSGVVDVQBFMIAPIGAETTFSEALRN 180
Qy 181 GAEVYHALKSVIKEGLSTGLGDEGGFAPSVGSTREALDLIVKAEKAGTTPGKDIALAL 240
Db 181 GAEVYHALKSVIKEGLSTGLGDEGGFAPSVGSTREALDLIVKAEKAGTTPGKDIALAL 240
Qy 241 DVASSEPFKDGTYHFEGGQSAEMANVYAEVLDAYPIVSIEDPLQEDDWEYTNLTATI 300
Db 241 DVASSEPFKDGTYHFEGGQSAEMANVYAEVLDAYPIVSIEDPLQEDDWEYTNLTATI 300
Qy 301 GDKVQIVGDDFFVTNPERLKEGIAKKAANSILVKVNOIGTITETFDVADMAHRAGYTSM 360
Db 301 GDKVQIVGDDFFVTNPERLKEGIAKKAANSILVKVNOIGTITETFDVADMAHRAGYTSM 360
Qy 361 SHRSGETEDTTIADLAVALNCGQIKTGAPARSDRVAKYNOLLRIEQLLDGAGVYAGRSAP 420
Db 361 SHRSGETEDTTIADLAVALNCGQIKTGAPARSDRVAKYNOLLRIEQLLDGAGVYAGRSAP 420
Qy 421 PRFQG 425
Db 421 PRFQG 425

RESULT 6
US-10-781-014-72
; Sequence 72, Application US/10781014
; Publication No. US20040180408A1
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Kröger, Burkhard
; APPLICANT: Schröder, Hartwig
; APPLICANT: Zeider, Oskar
; APPLICANT: Haberhauer, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
; INVOLVED IN CARBON METABOLISM AND ENERGY

; TITLE OF INVENTION: PRODUCTION
; FILE REFERENCE: BGI-126CPN
; CURRENT APPLICATION NUMBER: US/10/781.014
; CURRENT FILING DATE: 2004-02-17
; PRIOR APPLICATION NUMBER: US 09/602,740
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 60/141,031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: 60/143,208
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 60/151,572
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: DE 19931412.8
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931413.6
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931419.5
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931420.9
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931424.1
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931428.4
; PRIOR FILING DATE: 1999-07-08
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 784
; SEQ ID NO 72
; LENGTH: 425
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-10-781-014-72

Query Match 99.5%; Score 2145; DB 16; Length 425;
Best Local Similarity 99.5%; Pred. No. 4.6e-170;
Matches 423; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VAEIMHVFAREILDSRGNTVEAEVFLDDGSHGVAGVPSGASTGVHEAHLRGGDRYLG 60
DB 1 VAEIMHVFAREILDSRGNTVEAEVFLDDGSHGVAGVPSGASTGVHEAHLRGGDRYLG 60

QY 61 KGVLKAVENVNEIGDELADDDQRLIDEAMIKLDGTANKSRLGANAILGVSMVAKA 120
DB 61 KGVLKAVENVNEIGDELADDDQRLIDEAMIKLDGTANKSRLGANAILGVSMVAKA 120

QY 121 AADSAGLPLFRYTGGNNAHVLPVPMNIIINGGAHDSGVVDQBFMTAPIGAETFSALRN 180
DB 121 AADSAGLPLFRYTGGNNAHVLPVPMNIIINGGAHDSGVVDQBFMTAPIGAETFSALRN 180

QY 181 GAEVYHALKSVIKEKGLSTGLGDEGGFAPSVGSTREALDLIVKAIKAGFTPGKDIALAL 240
DB 181 GAEVYHALKSVIKEKGLSTGLGDEGGFAPSVGSTREALDLIVKAIKAGFTPGKDIALAL 240

QY 241 DVASSEFFKDGTVHFEGGQSHAAEMANVYAEVDAIPVSIEDPLQEDDWEGTNLTATI 300
DB 241 DVASSEFFKDGTVHFEGGQSHAAEMANVYAEVDAIPVSIEDPLQEDDWEGTNLTATI 300

QY 301 GDKVQIVGDDFFVTNPERLKEGIAKKAANSILVKVNIQIGLTETTFDAVDMHRAGYTSM 360
DB 301 GDKVQIVGDDFFVTNPERLKEGIAKKAANSILVKVNIQIGLTETTFDAVDMHRAGYTSM 360

QY 361 SHRSGETEDTTIADLAVALNCGQIKTGAPARSDRVAKYNOLLRIEQLLGDAGVYAGRSF 420
DB 361 SHRSGETEDTTIADLAVALNCGQIKTGAPARSDRVAKYNOLLRIEQLLGDAGVYAGRSF 420

QY 421 PRFQG 425
DB 421 PRFQG 425

RESULT 7
US-10-494-836-8
; Sequence 8, Application US/10494836
; Publication No. US20050014233A1

; GENERAL INFORMATION:
; APPLICANT: Zelder, Oskar
; APPLICANT: Pompejus, Markus
; APPLICANT: Schroder, Hartwig
; APPLICANT: Kroger, Burkhard
; APPLICANT: Klopptogge, Corinna
; APPLICANT: Haberhauer, Gregor
; TITLE OF INVENTION: Genes coding for proteins of carbon metabolism and energy products
; FILE REFERENCE: BGI-167US
; CURRENT APPLICATION NUMBER: US/10/494,836
; CURRENT FILING DATE: 2004-05-05
; PRIOR APPLICATION NUMBER: PCT/EP02/12135
; PRIOR FILING DATE: 2002-10-31
; PRIOR APPLICATION NUMBER: DE 101 54 270.4
; PRIOR FILING DATE: 2001-11-05
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 8
; LENGTH: 425
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-10-494-836-8

Query Match 99.5%; Score 2145; DB 17; Length 425;
Best Local Similarity 99.5%; Pred. No. 4.6e-170;
Matches 423; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VAEIMHVFAREILDSRGNTVEAEVFLDDGSHGVAGVPSGASTGVHEAHLRGGDRYLG 60
DB 1 VAEIMHVFAREILDSRGNTVEAEVFLDDGSHGVAGVPSGASTGVHEAHLRGGDRYLG 60

QY 61 KGVLKAVENVNEIGDELADDDQRLIDEAMIKLDGTANKSRLGANAILGVSMVAKA 120
DB 61 KGVLKAVENVNEIGDELADDDQRLIDEAMIKLDGTANKSRLGANAILGVSMVAKA 120

QY 121 AADSAGLPLFRYTGGNNAHVLPVPMNIIINGGAHDSGVVDQBFMTAPIGAETFSALRN 180
DB 121 AADSAGLPLFRYTGGNNAHVLPVPMNIIINGGAHDSGVVDQBFMTAPIGAETFSALRN 180

QY 181 GAEVYHALKSVIKEKGLSTGLGDEGGFAPSVGSTREALDLIVKAIKAGFTPGKDIALAL 240
DB 181 GAEVYHALKSVIKEKGLSTGLGDEGGFAPSVGSTREALDLIVKAIKAGFTPGKDIALAL 240

QY 241 DVASSEFFKDGTVHFEGGQSHAAEMANVYAEVDAIPVSIEDPLQEDDWEGTNLTATI 300
DB 241 DVASSEFFKDGTVHFEGGQSHAAEMANVYAEVDAIPVSIEDPLQEDDWEGTNLTATI 300

QY 301 GDKVQIVGDDFFVTNPERLKEGIAKKAANSILVKVNIQIGLTETTFDAVDMHRAGYTSM 360
DB 301 GDKVQIVGDDFFVTNPERLKEGIAKKAANSILVKVNIQIGLTETTFDAVDMHRAGYTSM 360

QY 361 SHRSGETEDTTIADLAVALNCGQIKTGAPARSDRVAKYNOLLRIEQLLGDAGVYAGRSF 420
DB 361 SHRSGETEDTTIADLAVALNCGQIKTGAPARSDRVAKYNOLLRIEQLLGDAGVYAGRSF 420

QY 421 PRFQG 425
DB 421 PRFQG 425

RESULT 8
US-10-282-122A-53908
; Sequence 53908, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangou
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant

Query Match	88.8%;	Score 1913;	DB 15;	Length 425;
Best Local Similarity	87.8%;	Pred. No.1e-150;		
Matches 373;	Conservative 23;	Mismatches 29;	Indels 0;	Gaps 0;
Qy 1	VAEIMHVAREIILDSRCNPTVEAEVFLDDGSHGVAGVPSGASTGVHEAHELDCGDRVLG	60		
Db 1	MADIMHVAREIILDSRCNPTVEAEVFLDDGSHGVAGVPSGASTGVHEAHELDCGGERVLG	60		
Qy 61	KGVLKAVENYNEETIGDELAGEADDQRLIDDEAMIKLDCGTANKSRIGANAILGVSMAVAKA	120		
Db 61	KGVLNAVNNVNEETIADAIAGAEADDQRLIDQAMTALDGTENKSRIGANAILGVSIYAKA	120		
Qy 121	AASAGLPLFRYIGGPNNAHLVPVPMNMIINGGAHADSGVDVQEFMIAPIGAETTFSEALRN	180		
Db 121	AAESAGLPLRYIGGPNNAHLVPVPMNMIINGGAHADSGVDVQEFMIAPIGAESFSEALRM	180		
Qy 181	GAEYTHALKSVIIEKGLSTGLDDEGGFAPSVGSTREALDLIVKAIKAGFTPGKDIALAL	240		
Db 181	GAEYTHSLKSVIIEKGLSTGLDDEGGFAPSVVESTKAALDLIVKAIKAGFPKPGADIALAL	240		
Qy 241	DVASSFFKDGTVHFEGGQHSAAEMANNVYAEFLVDAYPTVSTIEDPLQEDDWEGYTNLTATI	300		
Db 241	DVASSFFKDGTVHFEGGQHSAAEMANNVYAEFLVDAYPTVSTIEDPLQEDDWEGYTALTAAI	300		
Qy 301	GDVKQIVGDDFFVTNPNRLKEGIAKKAANSILVKVNOIGTITETFDVDMAHRAGYTSMW	360		
Db 301	GDVKQIVGDDFFVTNPNRLKEGIEKKAANALLVKVNOIGTITETFDVDMAHRYGRTMM	360		
Qy 361	SHRSGETEDTTIADLAVALNCGQIKTGAPARSDRVAKYNQLLRIEQLLGDAGVYAGRSAP	420		
Db 361	SHRSGETEDTTIADLAVALNCGQIKTGAPARSDRVAKYNQLLRIEQLLDDAAVYAGRSAP	420		
Qy 421	PRFQG 425			
Db 421	PRFQG 425			

Query Match	72.5%	Score 1562;	DB 15;	Length 423;
Best Local Similarity	72.8%;	Pred. No. 1.9e-121;		
Matches 308;	Conservative	50;	Mismatches 65;	Indels 0;
				Gaps 0;

Qy	1	VASIMHFAREILDSRGNPTVEAEVFLDDCGSHGVAGVPSGASTGVHAEHLRDCGGDRLYG	60
Db	1	LASIEAVHAREILDSRGNPTVEVEVALDDGTIARAAPSGASTGQFEAVELRDCGGDRLYG	60
Qy	61	KGVLKAVENNEIEGDELAGEADDDORLIDDEAMIKLDGTANKRGLGANAILGVSMVAVAKA	120
Db	61	KGVEKAVAAVNEIEISQIVGYEAEORLVDASLITLDGTPKSRGLGANAILGVSLAVAKA	120
Qy	121	AADSAGILPLFRYITGGPNNAHVLVPPVMNMIINGGAHADSGVDVQRFMIAPICGAETFEALRN	180
Db	121	AAESADLPLFRYITGGPNNAHVLVPPVMNMIINGGAHADTNVDIQEFMIAPICAESFEALRW	180
Qy	181	GA5VYHALKSVIKKGLSTGLGDEGGFAPSVGSTREALDLIVKAEKAGPTPGKDIALAL	240
Db	181	GA5VYHSLKAVLKAHGLATGVGDEGGFAPNLDNSRAALDLISEAITKAGFTLGGDIALAL	240
Qy	241	DVASSRFFKDGTYHFEQGQSHSAEMANVYAEILDVAYPIVSIEDPLQEDDDMEGYTNLTATI	300
Db	241	DVAATFEYADCAQFEGSSRSASEMAAYTELVESPLVSIEDPLSEEDWAGWKALTESL	300
Qy	301	GDKVQIVGDDFFVFTNPERLKEGIAKKAANSILVKVNIQIGTLTFTFDVADWAHRAGYTSMM	360
Db	301	GDRLQLVGDDLFVFTNPERLQRIQDEGAANSLLVKVNQIGTLTSETDLDAVSIAQRNRYTAMI	360
Qy	361	SHRSGETEDTTIADLAVALNCGOIKTCAPARSORVAKYNOLLRTTEOLLGAGVYAGRSFAF	420
Db	361	SHRSGETEDTTIADIADVATNSGQIKTCAPARSERVAKYNOLLRIEEELDDAAVTAGANAF	420
Qy	421	PRF 423	
Db	421	PRF 423	

```

RESULT 10
US-10-282-122A-62133
; Sequence 62133, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Heselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel

```

```

; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 62133
; LENGTH: 429
; TYPE: PRT
; ORGANISM: Mycobacterium avium
US-10-282-122A-62133

```

```

Query Match 72.4%; Score 1560.5; DB 15; Length 429;
Best Local Similarity 72.9%; Pred. No. 2.6e-121;
Matches 307; Conservative 47; Mismatches 66; Indels 1; Gaps 1;

QY 4 IMHVFAREILDSRGNTPTVEAEVFLDDGSHGVAGVPSGASTGVHEAHELRDGGDRYLKGV 63
Db 4 IEQVGAREILDSRGNTPTVEAEVFLDDGSHGVAGVPSGASTGVHEAHELRDGGDRYLKGV 63
QY 64 LKAVENNEEIGDELAGEADDQRLIDEAMIKLDGTANKSRILGANAILGVSMVAKAAAD 123
Db 64 QKAVQAVLDEIGPAVIGLNADDQRLVDQALVDLDTGTPDKSRILGANAILGVSLAVAKAAAD 123
QY 124 SAGLPLFRYIGGPNNAHLVPPMMNIIINGGAHDSGVVDVQEFMTAPIGAETFSALRWGAE 183
Db 124 SAGLPLFRYIGGPNNAHLVPPMMNIIINGGAHDSGVVDVQEFMTAPIGAETFSALRWGAE 183
QY 184 VYHALKSVIKKGLSTGLDGGGFPAPSVGSTRBALDLIVKAIKAGFTPGKDIALLDVA 243
Db 184 VYHLSKSVLKKEGLSTGLDGGGFPAPSVGSTRBALDLIVKAIKAGFTPGKDIALLDVA 243
QY 244 SSEFFKDGKT-YHFEQGGHSAEMANVYAEVLVDAYPIVSIEDPLQEDDWEQYTNLTATIGD 302
Db 244 ATEFYSDGTGYKEGSTRTEAQWAEYAGLLGAYPIVSIEDPLQEDDWEQYTNLTATIGD 303
QY 303 KVOIVGDDPFTVTPNERKEGIAKKAANSILVKVNIQTGLTETPDVADMAHRAGYTSMHSH 362
Db 304 RVQLVGDDVFTVTPNERKEGIEKGVANALLVKVNIQTGLTETLDAVALAHSGYRTMWSH 363
QY 363 RSETEDTTADIADVNCQIKTGAPASDRVAKYNQILRLTEQLLGDAGVAGSAPFR 422
Db 364 RSETEDTTADIADVAGSGQIKTGAPASDRVAKYNQILRLTEQLLGDAGVAGSAPFR 423
QY 423 F 423
Db 424 F 424

```

```

RESULT 11
US-10-282-122A-64490
; Sequence 64490, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 64490
; LENGTH: 429
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-10-282-122A-64490

```

```

Query Match 72.1%; Score 1553.5; DB 15; Length 429;
Best Local Similarity 73.2%; Pred. No. 9.8e-121;
Matches 308; Conservative 42; Mismatches 70; Indels 1; Gaps 1;

QY 4 IMHVFAREILDSRGNTPTVEAEVFLDDGSHGVAGVPSGASTGVHEAHELRDGGDRYLKGV 63
Db 4 IEQVGAREILDSRGNTPTVEAEVFLDDGSHGVAGVPSGASTGVHEAHELRDGGDRYLKGV 63
QY 64 LKAVENNEEIGDELAGEADDQRLIDEAMIKLDGTANKSRILGANAILGVSMVAKAAAD 123
Db 64 QKAVQAVLDEIGPAVIGLNADDQRLVDQALVDLDTGTPDKSRILGANAILGVSLAVAKAAAD 123
QY 124 SAGLPLFRYIGGPNNAHLVPPMMNIIINGGAHDSGVVDVQEFMTAPIGAETFSALRWGAE 183
Db 124 SAGLPLFRYIGGPNNAHLVPPMMNIIINGGAHDSGVVDVQEFMTAPIGAETFSALRWGAE 183
QY 184 VYHALKSVIKKGLSTGLDGGGFPAPSVGSTRBALDLIVKAIKAGFTPGKDIALLDVA 243
Db 184 VYHLSKSVLKKEGLSTGLDGGGFPAPSVGSTRBALDLIVKAIKAGFTPGKDIALLDVA 243
QY 244 SSEFFKDGKT-YHFEQGGHSAEMANVYAEVLVDAYPIVSIEDPLQEDDWEQYTNLTATIGD 302

```



```
Db 244 ATEFTDGTGVFEGTTRTADQMTFVAGLGGAYPLVSIEDPLSEDDWGDGWAALTASIGD 303
Qy 303 KVOIVGDDFFVTNPERLKEGIAKKAANSILVKVNOIGTFLTETPDVDMHRAAGVYTSMMSH 362
Db 304 RVQIVGDDIFVTNPERLEEGIERGVANALLVKVNOIGTFLTETPDVLAHGGVYRTMISH 363
Qy 363 RSGETEDTTIADLAVALNCGQIKTGAPARSDRVAKYNQLLRIEQLLGDAGVYAGRSAPPR 422
Db 364 RSGETEDTTIADLAVALNCGQIKTGAPARSDRVAKYNQLLRIEQLLGDAGVYAGRSAPPR 423
Qy 423 F 423
Db 424 F 424
```

RESULT 12

```
US-10-282-122A-62696
; Sequence 62696, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282.122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR FILING DATE: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 62696
; TYPE: PRT
; LENGTH: 439
; ORGANISM: Mycobacterium bovis
US-10-282-122A-62696
```

```
Query Match 72.1%; Score 1553.5; DB 15; Length 439;
Best Local Similarity 73.2%; Pred. No. 1e-120;
Matches 308; Conservative 42; Mismatches 70; Indels 1; Gaps 1;

Qy 4 IMHVAREILDSRGNPTVEAEVFLDDGSHGVAGVPSCGTCGVHAEHLRDCGDRYLKGV 63
Db 14 IEQVGAREILDSRGNPTVEAEVFLDDGSHGVAGVPSCGTCGVHAEHLRDCGDRYLKGV 73
```

```
Qy 64 LKAVENVNEBIGDELAGEADQRLIDEAMIKLDGTANKSRLGANAILGVSMVAKAAAD 123
Db 74 QKAVQAVLDEIGPAVIGLNADDQRLVDQALVDLDGTPDKSRLGNAILGVSLAVAKAAAD 133
Qy 124 SAGLPLFRYGGNAHVLPVPMNNIINGGAHAGSVVDVQEFMTAPIGAETTFSEALRNGAE 183
Db 134 SABLPLFRYVGGNAHILPVPMMNILLNGGAHADTAVDIOEFMVAPIGAPSFVEALRNGAE 193
Qy 184 VYHALKSVIKEKGLSTGLGDEGGFAPSVGSTRALDLI VKAIEKAGFTPGKDIALALDVA 243
Db 194 VYHALKSVLKEKGLSTGLGDEGGFAPDVAGTTAALDLISRAIESAGURPGADVALADAA 253
Qy 244 SSEFFKDDGT-YHFEFGQHSAAEMANYAEVLVDAYPIVSIEDPLQEDDWEGYTNLTATIGD 302
Db 254 ATEFTDGTGVFEGTTRTADQMTFVAGLGGAYPLVSIEDPLSEDDWGDGWAALTASIGD 313
Qy 303 KVOIVGDDFFVTNPERLKEGIAKKAANSILVKVNOIGTFLTETPDVDMHRAAGVYTSMMSH 362
Db 314 RVQIVGDDIFVTNPERLEEGIERGVANALLVKVNOIGTFLTETPDVLAHGGVYRTMISH 373
Qy 363 RSGETEDTTIADLAVALNCGQIKTGAPARSDRVAKYNQLLRIEQLLGDAGVYAGRSAPPR 422
Db 374 RSGETEDTTIADLAVALNCGQIKTGAPARSDRVAKYNQLLRIEQLLGDAGVYAGRSAPPR 433
Qy 423 F 423
Db 434 F 434
```

RESULT 13

```
US-10-156-761-11070
; Sequence 11070, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 11070
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-11070
```

```
Query Match 72.0%; Score 1551.5; DB 14; Length 428;
Best Local Similarity 72.9%; Pred. No. 1.4e-120;
Matches 310; Conservative 43; Mismatches 71; Indels 1; Gaps 1;
```

```
Qy 1 VAEIMEVFAREILDSRGNPTVEAEVFLDDGSHGVAGVPSCGTCGVHAEHLRDCGDRYL 59
Db 1 VPSIDVVVAREILDSRGNPTVEAEVFLDDGSHGVAGVPSCGTCGVHAEHLRDCGDRYL 60
Qy 60 GKGVLKAVENVNEBIGDELAGEADQRLIDEAMIKLDGTANKSRLGANAILGVSMVAK 119
Db 61 GKGVLKAVLAVIEQIGPELVGYDATEQRLIDQAMFOLDATDNKGLCANAILGVSLAVAH 120
Qy 120 AAADSAGLPLFRYGGNAHVLPVPMNNIINGGAHAGSVVDVQEFMTAPIGAETTFSEALR 179
Db 121 AAASEADPLFRYIGGNAHLLPVPMMNILLNGGSHADSNVDIOEFMTAPIGAESFSEALR 180
```

QY 180 NGAEVYHALKSVIKGLSTGLDGGGFPAPSVGSTREALDLIVKAIKAGFTPGKDIALA 239
 Db 181 WGAEVYHTLKKVLTGKSTGLDGGGFPAPLESNRAALDLIIBAIKQAGYIFGEQYALA 240
 QY 240 LDVASSEFFKDGTYHFGGQHSAAEMANVYAEVLVDAYPIVSIEDPLQEDDWEQYTNLTAT 299
 Db 241 LDVAASEFFKDGTYHFGGQHSAAEMTEYYEELVSAVPLVSIEDPLVEDDWDAGWKVITDK 300
 QY 300 IGDVKQIVGDDFFVTNPERLKEGIAKKAANSILVKVNIQIGTLTETFDVDMHAGVYTSM 359
 Db 301 LGDKVQIVGDDFFVTNPERLARGIEGSAEALLVKVNIQIGSLTETLDVAVLAQRNGPKCM 360
 QY 360 MSHRSGTETDTTADLAVNCGQIKTGAPARSDRVAKYNQILRIEQLLDGAGVYAGRSA 419
 Db 361 MSHRSGTETDTTADLAVNCGQIKTGAPARSDRVAKYNQILRIEQLLDAAEYAGRSA 420
 QY 420 FPRFQ 424
 Db 421 FPRF 425

RESULT 14
 US-10-282-122A-63682
 ; Sequence 63682, Application US/10282122A
 ; Publication No. US20040029129A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Liangsu
 ; APPLICANT: Zamudio, Carlos
 ; APPLICANT: Malone, Cheryl
 ; APPLICANT: Haselbeck, Robert
 ; APPLICANT: Ohlsen, Kari
 ; APPLICANT: Zykand, Judith
 ; APPLICANT: Wall, Daniel
 ; APPLICANT: Trawick, John
 ; APPLICANT: Carr, Grant
 ; APPLICANT: Yamamoto, Robert
 ; APPLICANT: Forsyth, R.
 ; APPLICANT: Xu, H.
 ; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
 ; FILE REFERENCE: EUTRA.034A
 ; CURRENT APPLICATION NUMBER: US/10/282,122A
 ; CURRENT FILING DATE: 2003-02-20
 ; PRIOR FILING DATE: 60/191,078
 ; PRIOR FILING DATE: 2000-03-21
 ; PRIOR APPLICATION NUMBER: 60/206,848
 ; PRIOR FILING DATE: 2000-05-23
 ; PRIOR APPLICATION NUMBER: 60/207,727
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: 60/230,335
 ; PRIOR FILING DATE: 2000-09-06
 ; PRIOR APPLICATION NUMBER: 60/230,347
 ; PRIOR FILING DATE: 2000-09-09
 ; PRIOR APPLICATION NUMBER: 60/242,578
 ; PRIOR FILING DATE: 2000-10-23
 ; PRIOR APPLICATION NUMBER: 60/253,625
 ; PRIOR FILING DATE: 2000-11-27
 ; PRIOR APPLICATION NUMBER: 60/257,931
 ; PRIOR FILING DATE: 2000-12-22
 ; PRIOR APPLICATION NUMBER: 60/267,636
 ; PRIOR FILING DATE: 2001-02-09
 ; PRIOR APPLICATION NUMBER: 60/269,308
 ; PRIOR FILING DATE: 2001-02-16
 ; Remaining prior application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 78614
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 63682
 ; LENGTH: 447
 ; TYPE: PRT
 ; ORGANISM: Mycobacterium leprae
 US-10-282-122A-63682

Query Match 69.7%; Score 1502.5; DB 15; Length 447;
 Best Local Similarity 70.8%; Pred. No. 1.9e-116;

Matches 300; Conservative 47; Mismatches 76; Indels 1; Gaps 1;
 QY 1 VAEIMHVFAREILDSRGNTVEAEVFLDDGSHGVAGVPSGASTGVVHEAHELDDGDRYLG 60
 Db 19 VPVIEQVGAREILDSRGNTVEEVVLIDGTTFARAAVPSGASTGEVEAEVLDDGDRYGG 78
 QY 61 KGVLKAVENVNBEIGDELADDDORLIDEAMIKLDGTANKSRILGANAILGVSMVAKA 120
 Db 79 KGVKRAVDVLDIBGPVWITGLANDORLIDQELLDDGTDPKSRILGNAILGVSLAVAKA 138
 QY 121 AADSAGLPLFRYIGGNNAHVLPVPMNIIINGGAHDSGVVDQEFMTAPIGAETFSBALRN 180
 Db 139 AADSAGLPLFRYIGGNNAHVLPVPMNIIINGGAHDSGVVDQEFMTAPIGAETFSBALRN 198
 QY 181 GAEVYHALKSVIKGLSTGLDGGGFPAPSVGSTREALDLIVKAIKAGFTPGKDIALAL 240
 Db 199 GAEVYHALKSVIKGLSTGLDGGGFPAPSVGSTREALDLIVKAIKAGFTPGKDIALAL 258
 QY 241 DVASSEFFKDGTYHFGGQHSAAEMANVYAEVLVDAYPIVSIEDPLQEDDWEQYTNLTAT 299
 Db 259 DAAATEFYTDGTYHFGGQHSAAEMANVYAEVLVDAYPIVSIEDPLQEDDWEQYTNLTAT 318
 QY 300 IGDVKQIVGDDFFVTNPERLKEGIAKKAANSILVKVNIQIGTLTETFDVDMHAGVYTSM 359
 Db 319 IGEQVQIVGDDFFVTNPERLKEGIAKKAANSILVKVNIQIGTLTETFDVDMHAGVYTSM 378
 QY 360 MSHRSGTETDTTADLAVNCGQIKTGAPARSDRVAKYNQILRIEQLLDGAGVYAGRSA 419
 Db 379 ISHRSGTETDTTADLAVNCGQIKTGAPARSDRVAKYNQILRIEQLLDGAGVYAGRSA 438
 QY 420 FPRF 423
 Db 439 FURY 442

RESULT 15
 US-10-369-493-11696
 ; Sequence 11696, Application US/10369493
 ; Publication No. US2003033675A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Chen, Xianfeng
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
 ; FILE REFERENCE: 38-10(52052)B
 ; CURRENT APPLICATION NUMBER: US/10/369,493
 ; CURRENT FILING DATE: 2003-02-28
 ; PRIOR APPLICATION NUMBER: US 60/360,039
 ; PRIOR FILING DATE: 2002-02-21
 ; NUMBER OF SEQ ID NOS: 47374
 ; SEQ ID NO 11696
 ; LENGTH: 423
 ; TYPE: PRT
 ; ORGANISM: Agrobacterium tumefaciens
 US-10-369-493-11696

Query Match 66.0%; Score 1422; DB 15; Length 423;
 Best Local Similarity 66.2%; Pred. No. 8.9e-110;
 Matches 276; Conservative 58; Mismatches 81; Indels 2; Gaps 1;
 QY 4 IMHVFAREILDSRGNTVEAEVFLDDGSHGVAGVPSGASTGVVHEAHELDDGDRYLGKV 63
 Db 4 ITDIIAREILDSRGNTVEVDVLYLEDGSMGRAVPSGASTGAEVAVELDDGDRYLGKV 63
 QY 64 LKAVENVNBEIGDELADDDORLIDEAMIKLDGTANKSRILGANAILGVSMVAKAAD 123
 Db 64 EKAVEAVNTEIFDAIGGFAENQIQDMMIALDGTANKSRILGANAILGVSMVAKAAD 123
 QY 124 SAGLPLFRYIGGNNAHVLPVPMNIIINGGAHDSGVVDQEFMTAPIGAETFSBALRN 183

Db	124	ASGLPLYRYVGGPNAHLPPVPMNIIINGGAHADNPIDFQEFMILPVGAENIREAVRMGSE	183
Qy	184	VYHALKSVIKEKGLSTGLGEGGFAPSVGSGTREALDLIVKAIKAGETPCGKDIALALDVA	243
Db	184	VFTLKKELSAQHNTNVGDEGGFAPGLESAPALDFIMKSIKAGYRPGEDMYVGLDCA	243
Qy	244	SSEFFKDGTYHFEQGHS--AAEMANYAELVDAYPIVSIEDPLQEDDWEQYTNLTATIG	301
Db	244	STEFFKDGKYLEGEGRTEPGAMAEYLAELVNKYPIISVEDGMAEDDWEQWKTLTDLVG	303
Qy	302	DKVQIVCDDDFVTNPERLKEGIAKKAANSILVKVNOIGTLTETFDVDAVMAHRAGYTSMS	361
Db	304	NKCOLVGDDLFVTNSARLRDGIKXGVANSILVKVNOIGSLSETLDAVETAHKAGYTAVMS	363
Qy	362	HRSGETEDTTIADLAVALNCGQIKTGAPARSDRVAKYNQLLRIEQLIGDAGVYAGRS	418
Db	364	HRSGETEDSTIADLAVATNCGQIKTGSLSARSDRLAKYNQLIRIEEMLGPQAYAGRS	420

Search completed: September 29, 2005, 14:45:40
Job time : 1027 secs

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)